THE AMERICAN MATHEMATICAL MONTHLY

(FOUNDED IN 1894 BY BENJAMIN F. FINKEL)

THE OFFICIAL JOURNAL OF

THE MATHEMATICAL ASSOCIATION OF AMERICA, INC.

VOLUME 78 NUMBER 2

CONTENTS

(*Continued on inside cover*)

FEBRUARY 1971

NOTICE TO AUTHORS

Specialized research is usually unsuitable; see *Statement of Policy* (vol. 76, p. 2). Manuscript preparation: Please use the *Manual for Monthly Authors* (vol. 78, p. 1) and follow the format in current issues of the MONTHLY. Manuscripts should be typewritten, triple-spaced with wide margins; submit two copies and keep one for protection against loss. Each author receives 50 free reprints.

Backlog: Main Articles 8 months, Math. Notes 8 months, Research Problems 6 months, Classroom Notes 7 months, Math. Education 6 months.

EDITORIAL CORRESPONDENCE AND MAIN ARTICLES: to HARLEY FLANDERS, American Mathematical Monthly, Tel Aviv University, Ramat Aviv, Israel (see Notice, Vol. 77, p. 555); NOTES, etc. to the corresponding Associate Editor; ADVERTISING CORRESPONDENCE: to RAOUL HAILPERN, Mathematical Association of America, SUNY at Buffalo, Buffalo, N. Y. 14214; CHANGE OF ADDRESS and SUBSCRIPTIONS: to A. B. WILLCOX, Mathematical Association of America, 1225 Connecticut Ave., N.W., Washington, D. C. 20036

Annual dues for members of the Association (including a subscription to the American Mathematical Monthly) are $12.50. For nonmembers the subscription price is $18.00.

PUBLISHED BY THE ASSOCIATION at Washington, D. C., and Menasha, Wisconsin during the months of January, February, March, April, May, June-July, August-September, October, November, December.

Second-class postage paid at Washington, D. C., and additional mailing officies. Acceptance for mailing at special rate of postage provided for in the Act of February 28, 1925, embodied in Paragraph 4, Section 538, P. L. and R., authorized April 1, 1926.

PRINTED IN THE UNITED STATES OF AMERICA

BURTON W. JONES

AWARD FOR DISTINGUISHED SERVICE TO PROFESSOR BURTON WADSWORTH JONES

This year's recipient of the Mathematical Association of America's Award for Distinguished Service, Burton W. Jones, has in a quiet but very effective way made remarkable contributions to the cause of mathematics education at three levels: local, national, and international. Born in Minnesota, he did his undergraduate work at Grinnell College, went to Harvard for an M.A., and received his Ph.D. from the University of Chicago in 1928. He taught briefly at Western Reserve and the University of Chicago before settling down for a while at Cornell where he was promoted to a full professorship in 1945.

In 1948 he was persuaded to move West to become Chairman of the Department of Mathematics at the University of Colorado. Here he was responsible not only for greatly strengthening the department but also for making it, at the same time, an internationally known center for studies in number theory. He relinquished the chairmanship in 1963 but has continued to contribute to the welfare of the department by taking on short term but crucial administrative tasks from time to time.

Still at the local level, but in a broader context, Professor Jones was the initiating force in the creation of an informal regional organization which brings together faculty members from universities, four-year colleges and junior colleges in Colorado and Wyoming to discuss mutual problems in mathematics education. He has also taken an active interest in school mathematics and has given much time to lectures and informal discussions with high school students and teachers in his area.

At the national level, Burton Jones has made contributions to mathematics education through a wide variety of channels. The list of offices he has held and committees he has served on is too long to be presented now in full, and only a sample will be mentioned. He has been very active in the affairs of this Association. He was the first Associate Secretary, serving from 1943 to 1947; he has served on the Board of Governors, and has been a vice-president and a representative of the Association on the Council of the American Association for the Advancement of Science. He has served on numerous Association committees, often as chairman.

He has also taken part in the activities of many other organizations, including the American Mathematical Society and the National Council of Teachers of Mathematics. In the latter, he served two terms on its Board of Directors, and has carried out many committee assignments. In addition, he has been quite active in the School Mathematics Study Group. He was a member of the writing team which prepared the original junior high school volumes, assisted in the classroom tryout of early versions of SMSG texts, served on various advisory committees, and wrote a book explicitly for junior high school teachers at the request of SMSG. In this connection, it should be mentioned that without any outside urging, he has written several other books useful in mathematics education at both the pre-college and the college level.

One of Burton Jones' most important committee assignments was to the Committee on Regional Development of Mathematics of the Division of Mathematics of the National Research Council. He participated in three recommendations made by this Committee in 1952, each of which was to have profound effects on mathematics education in this country. The first was a recommendation to the National Science Foundation that it establish a program of summer institutes. The second, again to the National Science Foundation, was a recommendation that a program of visiting lecturers be established. The third was a recommendation to this Association that a reform of the undergraduate program in mathematics be undertaken.

It should be observed that Professor Jones backed up these recommendations by participating in the visiting lecturer program after it was set up, and by directing the first two summer institutes in mathematics funded by NSF.

At the international level, most of his activities have been in Latin America. He spent a full year in Central America working with the mathematics departments in the five national universities there in their attempts to improve their undergraduate programs. He has made a number of other visits to Central America and to various points in South America. Thanks to his patience and tact he accomplished much, and everywhere he went he left behind feelings of respect, fondness, and gratitude for his work.

But the same is true about him here at home. And the respect, affection, and appreciation he receives is indeed well earned. His efforts to improve mathematics education have been not only numerous but also substantially successful.

One of the reasons for his success is that he is very firm, almost stubborn. When he makes up his mind as to what is right, it is usually impossible to get him to change. When he decides something needs to be done, it gets done.

But he is at the same time a very gentle man. While he holds his beliefs strongly, he imposes them on no one. He will provide, at the appropriate time, an occasional quiet suggestion, but never a command.

His own father once put it in slightly different words when he remarked, "One thing you can say about Burton: he does not have many rough edges." That was said many years ago, but it is still true, and the mathematical community is the better for it.

E. G. BEGLE

AWARD OF THE 1971 CHAUVENET PRIZE TO PROFESSOR NORMAN LEVINSON

The Board of Governors of the Mathematical Association of America at its meeting on August 23, 1970, at the University of Wyoming voted to award the 1971 Chauvenet Prize to Professor Norman Levinson of the Massachusetts Institute of Technology for his paper "A Motivated Account of an Elementary

Proof of the Prime Number Theorem," published in this MONTHLY, 76 (1969), 225–245.

A certificate and monetary award in the amount of five hundred dollars was presented to Professor Levinson at the time of the Business Meeting of the Association on January 24, 1971, in Atlantic City.

The Chauvenet Prize is awarded for a noteworthy paper of an expository or survey nature published in English, which comes within the range of profitable reading for members of the Association. The purpose of the prize is to stimulate the writing of expository and survey articles. The 1971 Prize, awarded for a paper published in the three-year period 1967–69, is the nineteenth award of the Chauvenet Prize since its institution by the MAA in 1925. For a list of the names of previous winners, see this MONTHLY, 71 (1964), p. 589, 72 (1965), pp. 2–3, 74 (1967), p. 3, 75 (1968), pp. 3–4, and 77 (1970), pp. 117–118.

Professor Levinson was born on August 11, 1912 in Lynn, Massachusetts. He received the B.S. and M.S. degrees in electrical engineering in 1934 from the Massachusetts Institute of Technology. While a student of electrical engineering he also studied mathematics intensively and essentially completed the requirements for the degree D.Sc. in Mathematics with Norbert Wiener. Wiener obtained an M.I.T. travelling fellowship for him to spend 1934–35 with G. H. Hardy in Cambridge, England. (Later D. C. Spencer, also an engineering student, was awarded the same M.I.T. fellowship to Cambridge, where he studied with Littlewood.) In 1935 Levinson was awarded the D.Sc. by M.I.T. For the next year and one-half he was a National Research Council Fellow at Princeton University and the Institute for Advanced Study, where he continued his research in gap and density theorems under the nominal supervision of von Neumann. In February 1937 he became an instructor at M.I.T., where he has been except for 1948–49 when he was a Guggenheim Fellow at the Mathematics Institute in Copenhagen and 1967–68 when he was at the University of Tel Aviv.

After the publication of his Colloquium book, *Gap and Density Theorems*, by the American Mathematical Society in 1940, he shifted his research to differential equations for which he was awarded the Bôcher Prize by the A.M.S. in 1953. He was a Vice President of the A.M.S. in 1966–68. In 1955, *Theory of Ordinary Differential Equations* by Earl Coddington and Levinson was published. Since then he has also published articles in probability, complex linear programming, and analytic number theory.

He is a member of the National Academy of Science and the American Academy of Arts and Sciences.

He has supervised the theses of some 34 Ph.D. students. He is presently Chairman of the M.I.T. Mathematics Department.

In accepting the Award, Professor Levinson expressed himself as very pleased that the Association had honored him with the 1971 Chauvenet Prize. He was particularly gratified when he found that his old teacher G. H. Hardy had been an early recipient of the Prize.

THE KAKEYA PROBLEM FOR SIMPLY CONNECTED AND FOR STAR-SHAPED SETS

F. CUNNINGHAM, JR., Bryn Mawr College

A plane set K is called a **Kakeya set** if a unit segment, hereafter referred to as the **needle,** can be continuously turned around in K so as to return to its original position with its ends reversed. The problem posed by Kakeya was to find a Kakeya set of minimum area. He apparently thought that the answer was the deltoid (three-cusped hypocycloid) with area $\pi/8$. Besicovitch surprised everybody by showing [2, 3] that actually there exist Kakeya sets of arbitrarily small area. His beautiful solution, with improvements of exposition supplied by Perron and Schoenberg, has become widely known through his film produced by the Mathematical Association of America.

The familiar examples of Kakeya sets of small area are highly multiply connected and have large diameters. This is because these figures use "Pál joins" as a device for shifting the needle parallel to itself using only a small area. Each Pál join consists of a long excursion along one line and return along another line almost parallel to the first, the transition from one line to the other being made by a small triangle where they meet. Except for that triangle the area enclosed between the two lines is not used and is not to be counted.

It was therefore natural to ask, and Besicovitch underlined this problem, whether the area of a Kakeya set can still be arbitrarily small when the set is required to be simply connected. This is one of a cluster of similar problems, namely the Kakeya problem restricted to sets of various specified classes. Two such problems have been solved. Already in 1921 J. Pál [6] showed that the solution of Kakeya's problem for convex sets is the equilateral triangle of area $1/\sqrt{3}$. Then in a little-known paper of 1941 [1] A. H. van Alphen used some of the same ideas I shall use here to show that a Kakeya set can have arbitrarily small area and stay inside a circle of radius $2+\epsilon$ (arbitrary positive ϵ), thus eliminating the unboundedness of the Besicovitch examples.

Discounting some inconclusive but suggestive evidence by R. J. Walker [7] the first simply connected Kakeya sets with areas improving on the deltoid were formed in 1965. The best of these, found independently by Melvin Bloom and I. J. Schoenberg [4], have areas approaching what I shall call the **Bloom-Schoenberg number,** namely $(5-2\sqrt{2})\pi/24$, or approximately $.09\pi$ (to be compared with $\pi/8$ for the deltoid). Schoenberg conjectured that this number is the answer to the problem—that no simply connected Kakeya set can have an area any smaller than that. It turns out that he was wrong. The following result settles the question, at the same time improving on van Alphen.

Frederic Cunningham, Jr. received his Harvard PhD in 1953 under L. H. Loomis. He has held positions at the Univ. of New Hampshire, Wesleyan Univ., and Bryn Mawr College, and spent a sabbatical year at Orsay and Paris. His research is mostly on Banach spaces, and he received an M.A.A. Ford Award in 1967. *Editor.*

THEOREM 1. *Given $\epsilon>0$, there exists a simply connected Kakeya set of area less than ϵ contained in a circle of radius* 1.

In passing it may be remarked that the radius of the circle cannot be further reduced in Theorem 1. Indeed, if D is a disk of radius $r<1$, and if P is a point of D within a distance $1-r$ of the center, then $D-P$ is not a Kakeya set. (We are using a definition of Kakeya set which requires the needle to return to its starting position with reversed direction.) Hence any Kakeya set contained in D must contain the small disk of radius $1-r$ concentric with D, and must therefore have area at least $\pi(1-r)^2$. Simple connectedness is not at issue in this remark.

A class of sets intermediate between convex sets and simply connected sets is the class of star-shaped sets. A set K is called **star-shaped** with respect to O if the line segment joining O to any point of K is contained in K. The deltoid and the Bloom-Schoenberg examples are star-shaped, while the examples I shall construct to prove Theorem 1 are not. The Kakeya problem for star-shaped sets is not completely solved, but at least I can show that the situation here is different from that for simply connected sets.

THEOREM 2. *Every star-shaped Kakeya set has area at least* $\pi/108$.

We do not know if there are star-shaped Kakeya sets with areas between $\pi/108$ and the Bloom-Schoenberg number.

1. A Crude Beginning. To prove Theorem 1 we must make small simply connected Kakeya sets. Intuitively, a Kakeya set is made by turning a unit segment (the needle) continuously around, keeping track of the set of points touched by the needle in its motion. We can, of course, enlarge the set at will, for example by filling in holes to make it simply connected. The needle is free to move anywhere along the line it is in, with no cost of area. This fact is used in Besicovitch's solution to show that for an arbitrarily small cost in increased area, the needle can be moved from one line to any other line parallel to it. Once this is shown the problem becomes one of *turning* the needle—only its direction matters. The same is not true for proving Theorem 1. We always must be concerned with the *line* on which the needle lies, not only with the direction of that line.

The examples which prove Theorem 1 are going to be spidery sets in the following general format. In the center there will be a small convex nucleus Π from which radiate a large number of slender triangles in different directions. The needle turns a little in one triangle then slides lengthways through the nucleus to a triangle on the other side, there to turn a little more. The more detailed description of the examples is arrived at only by a lengthy succession of constructions, each replacing one Kakeya set by another of smaller area, while carefully preserving the simple connectedness. In these replacements only the radiating triangles are changed, not the nucleus.

The starting point is the following Kakeya set, shown in Figure 1. Let Γ be a fixed unit circle. When $\epsilon>0$ has been given, let Π be the region bounded by a

regular polygon concentric with Γ, having a large odd number Q of sides and so small that its area is less than $\frac{1}{2}\epsilon$. The vertices of Π are so numbered $C_0, C_1, \cdots, C_Q = C_0$ that successive vertices are almost opposite each other, the clockwise angle from C_{q-1} to C_q being $\pi Q/(Q+1)$ for each q in $\{1, \cdots, Q\}$. Now Q triangles $J_1, \cdots, J_Q$ are formed: J_q has C_q for one vertex, the other two vertices A_q and B_q being the points on Γ determined by extending $C_{q-1}C_q$ and $C_{q+1}C_q$ respectively. The set K of interest is the union of the polygonal region Π and the triangular regions $J_1, \cdots, J_Q$.

This set is obviously simply connected and contained inside Γ.

The needle is supposed to move in K from a segment $C_{q-1}C_qA_q$ to $C_{q+1}C_qB_q$ by turning around C_q as pivot. For this to be possible we need only to require that the perpendicular distance between the parallel lines $C_{q-1}C_{q+1}$ and A_qB_q is at least 1. This is achieved by making Q large enough. (Indeed, no matter how small ϵ is, as $Q \to \infty$, Π approaches a circle of area $\frac{1}{2}\epsilon$, hence of radius $\sqrt{\epsilon/2\pi}$, and so the distance in question tends to $1+\sqrt{\epsilon/2\pi} > 1$ as a limit.) When this condition is satisfied K is a Kakeya set. The needle turns around in K as follows: Starting, say, on the segment $C_0C_1A_1$ it turns around C_1 until it is on $C_2C_1B_1$; then it slides along that line until it is on $A_2C_2C_1$; then it turns around C_2 until it is on $B_2C_2C_3$; then it slides along that line until it is on $C_2C_3A_3$; and so on. After Q such turnings and slidings the indices have changed by Q, that is, it is back on $A_1C_1C_0$. It arrives there with its direction reversed, Q being odd.

The area of K is too big to be interesting. For large Q and small ϵ it is almost $\frac{1}{2}\pi$. To see how this can be improved, read on.

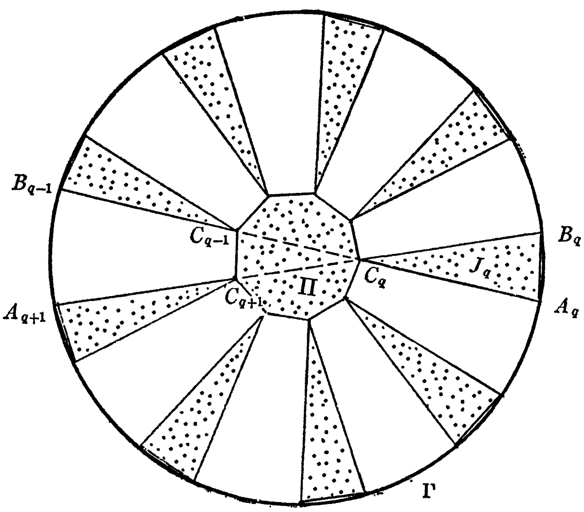


FIG. 1

2. The First Improvement. We focus attention on the problem of moving the needle from one line to another in a smaller area—in Fig. 1 from $C_{q-1}C_qA_q$

to $C_{q+1}\, C_q B_q$. If we can do this, then we can repeat it Q times to turn the needle all the way.

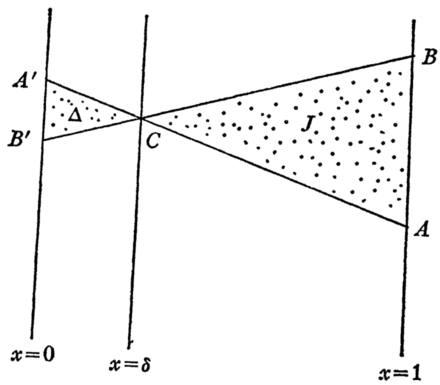


FIG. 2

In Figure 2 is displayed with simplified notation the part of Figure 1 relevant to this immediate purpose. Here C_{q-1} and C_{q+1} have been renamed A' and B' and all the subscripts involving q have been discarded. A coordinate system has been supplied as shown, the x-axis being omitted, because its position is immaterial. The smaller triangle Δ with vertices $A'B'C$ is part of the polygonal nucleus Π. Its altitude is called δ. The large triangle J with vertices ABC has altitude r, and for convenience we have made $\delta+r=1$, which is just large enough. Our object is to replace the obvious needle motion from $A'A$ to $B'B$ in $K^{(0)}=\Delta\cup J$ (namely turning around C) by a trickier one using less area. We shall keep Δ, which is acceptably small already, but get rid of most of J. The construction consists of a sequence of replacements, of which I am describing here only the first. It is shown in Figure 3.

The new figure consists of a **tree** (the part in the half-plane $x\geqq 0$), and a union of **joins** (the triangles in the half-plane $x\leqq 0$). In general terms, trees are made as follows. For some positive integer m choose M distinct points $C_1, \cdots, C_m$ in descending order, strictly between A' and B' on the y-axis, and $m-1$ distinct points in ascending order strictly between A and B on the line $x=1$. To complete the notation set $C_0=A'$, $C_{m+1}=B'$, $V_0=A$, $V_m=B$. The tree T is the union of the $m+1$ triangles t_i with vertices $C_iV_iC_{i+1}$ for $i=0, \cdots, m$.

The joins for T are the m triangles $J_1, \cdots, J_m$, where J_i has for vertices C_i and the points A_i and B_i on the vertical line $x=-r$ obtained by extending backward the segments C_iV_{i-1} and C_iV_i, which are sides of two adjacent triangles t_{i-1} and t_i of the tree.

The union $K^{(1)}$ of T and $J_1, \cdots, J_m$ will be called the **first replacement** of $K^{(0)}=\Delta\cup J$. Like $K^{(0)}$ it is a "partial Kakeya set," which means that in it the needle can move from $A'A$ to $B'B$. It does so as follows. Starting on $A'A=C_0V_0$ it turns in t_0 around V_0 until it is on C_1V_0; then it slides along that line until it

is between $x=-r$ and $x=\delta$; then it turns in $J_1\cup\Delta$ around C_1 until it is on $B_1C_1V_1$; then it slides along that line until it is between $x=0$ and $x=1$; then it turns around V_1 in t_1 until it is on C_2V_1; and so on until it arrives on $C_{m+1}V_m$, which is $B'B$.

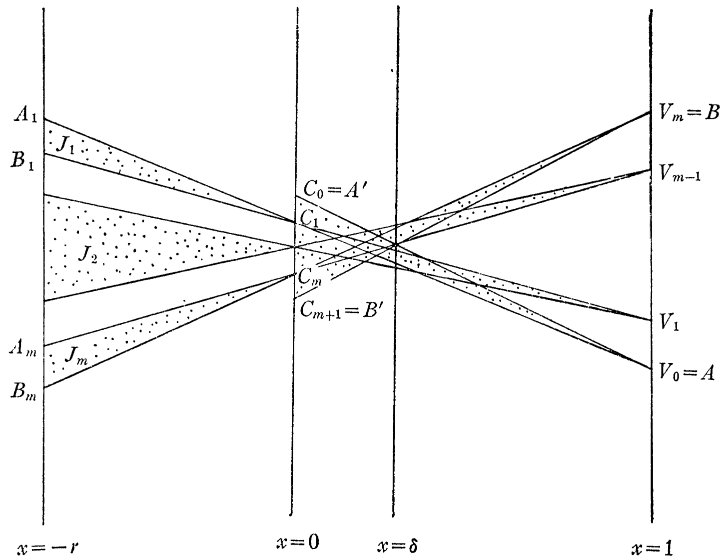


FIG. 3

We hope that $K^{(1)}$ is simply connected, and Figure 3 suggests that this is so. Postponing the proof for the tree, we can see that the joins do not spoil simple connectedness, because they do not intersect each other. Indeed the slope of the lower side of any join J_i is less than the slope of the upper side of its neighbor below, J_{i+1}, because these lines meet at V_i, which is to the right of $x=0$.

The purpose of the construction of $K^{(1)}$ was to save area. Has it succeeded? Let us compute first the total area of all the joins. Write a for the area of the original triangle J and a_i for the area of $J_i(i=1,\cdots,m)$. For each $i=1,\cdots,m$ we read from the similar triangles $J_i=A_iC_iB_i$ and $V_{i-1}C_iV_i$ the proportionality $|A_iB_i|=r|V_{i-1}V_i|$. (Absolute value signs as used here and elsewhere denote lengths of segments.) Adding these up we have $\sum|A_iB_i|=r|AB|$, whence

$$\sum a_i = \sum \tfrac{1}{2}r|A_iB_i| = \tfrac{1}{2}r^2|AB| = ra.$$

Conclusion: the contribution of the joins to the area of $K^{(1)}$ is slightly less than the contribution of J to the area of $K^{(0)}$. It remains to see how the area of T compares with the area of Δ, and this requires a more detailed specification of how T is made.

3. Trees. The success of the device described in the preceding section depends on two questions. First, can the tree be made to have small enough area

so that altogether $K^{(1)}$ has less area than $K^{(0)}$? Second, if the answer to that question is favorable, can the slight area improvement (not better than multiplication by $r=1-\delta$) be used as the beginning of a much more significant improvement? This section settles the first question. You are reminded that T contains Δ. We shall show that the area of T can be made to exceed that of Δ by as little as we please.

This is of course the crux of the solution of Kakeya's problem. It is here that Besicovitch made his decisive contribution, and it is his insight which we shall use in a modified form. The trees to be constructed here are akin to, but different from, the Besicovitch-Perron-Schoenberg trees found in [3], the difference being that in the B-P-S trees neighboring elementary triangles are almost parallel to each other, which makes the joins connecting them necessarily very far away, whereas our trees are explicitly designed so that neighboring elementary triangles are pointed towards joins with vertices right on the y-axis.

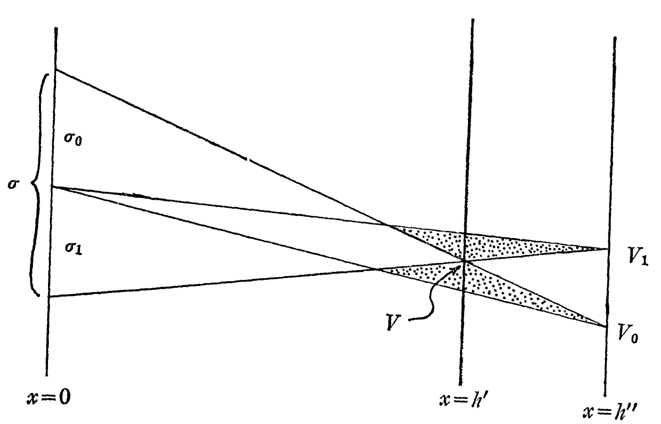


FIG. 4

For each positive integer p there is a special tree T_p which grows from the triangle Δ, called the **thin tree** of order p on Δ. It takes shape by repetition of an operation called **sprouting**, which transforms a triangle into the union of two triangles. Sprouting is illustrated in Figure 4. Let $0<h'<h''$, and let t be a triangle with its base σ on the y-axis and its vertex V on $x=h'$. Then sprouting from h' to h'' replaces t by $t_0 \cup t_1$, where t_0 and t_1 are the following triangles: their bases, on the y-axis, are the upper and lower halves σ_0 and σ_1 respectively of σ, and their remaining vertices are the points V_0 and V_1 on $x=h''$ found by extending the upper and lower sides, respectively, of t.

Now T_p is defined as follows. Divide the x-interval $[\delta, 1]$ into p equal parts by inserting points $h_1<h_2< \cdots <h_{p-1}$ between $\delta=h_0$ and $1=h_p$, so that $h_i-h_{i-1}=r/p$ for $i=1, \cdots, p$. Now sprout Δ from h_0 to h_1 to make two triangles, sprout both of these from h_1 to h_2 to make four triangles, and so on. The union

of the 2^p triangles you get from p successive sproutings is T_p. The case $p=4$ is shown in Figure 5.

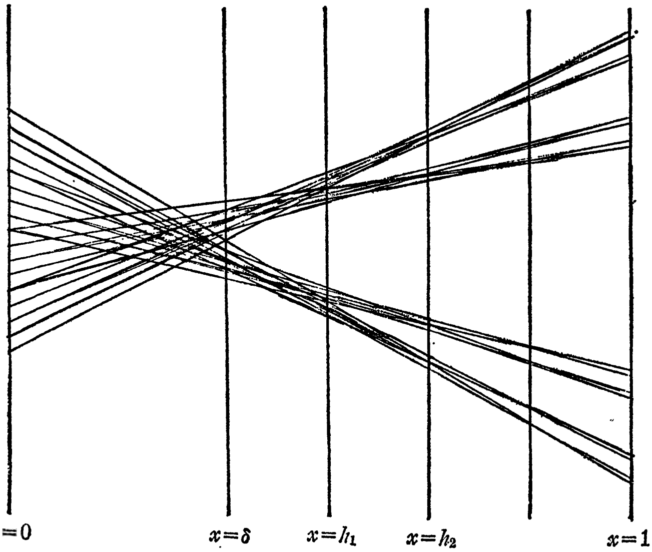


FIG. 5

To see that T_p is a special case of tree as defined in Section 2, we shall exploit a system of binary indices for the triangles which make it up. Each time a triangle sprouts, the two new triangles coming from it are named by affixing the subscripts 0 and 1 respectively to the notation for the triangle they replace, just as was done in defining sprouting. Starting with $t=\Delta$ (no subscripts), we get first t_0 and t_1, then t_{00}, t_{01}, t_{10}, and t_{11}, and so on. After p sproutings, each elementary triangle of T is represented by t with a string of p subscripts, which is the binary representation of one of the integers $0, 1, \cdots, m=2^p-1$. We can if we please read the string of subscripts as meaning that integer, and when we do so we observe (by induction on p) that the $m+1=2^p$ vertices $V_0, \cdots, V_{m+1}$ on $x=1$ of the triangles are numbered in ascending order as they should be, and that their bases $\sigma_0, \cdots, \sigma_m$ on the y-axis are the segments of the partition of $A'B'$ into 2^p equal parts numbered in descending order. We of course take the division points of this partition to be in descending order $C_1, \cdots, C_m$ and we have again Figure 3.

Finding the area of T_p is more complicated than for the B-P-S trees, and we shall content ourselves with an upper estimate. The first step is to estimate the shaded area in Figure 4, which is the increment of area resulting from a single sprouting. It is the union of two triangles which though not congruent have the same area. Both are contained in a vertical strip of horizontal width $2(h''-h')$, and each has vertical width at $x=h$, its widest point, which by similar triangles comes out to be $\frac{1}{2}|\sigma|(h''-h')/h''$. Therefore the shaded area is less than $|\sigma|(h''-h')^2/h''$.

Now consider T_p. We do not count Δ, which is part of $K^{(0)}$, but only the increment caused by growing T_p from Δ. At the kth stage of sprouting 2^{k-1} triangles simultaneously sprout to make 2^k triangles. For each of these σ is $2^{-(k-1)}|A'B'|$, $h''-h'$ is r/p and $h''>\delta$. By the formula from the last paragraph, all the sproutings of this stage contribute less than $2^{k-1}[2^{-(k-1)}|A'B'|(r/p)^2/\delta] = |A'B'|r^2/p^2\delta$ of new area. Adding the contributions from all p stages gives for the area of $T-\Delta$ less than $|A'B'|r^2/p\delta$. This tends to 0 as p becomes large. (We remark that $|A'B'|=|AB|\delta/r$, and that the area of J is $a=\frac{1}{2}r|AB|$, so that our estimate can be more neatly expressed as $2a/p$.)

A descriptive property of T_p for large p which will be useful later is that it is very thin in the waist. By the **waist** of T_p, I mean the segment of the vertical line $x=\delta$ in which T_p crosses it. The upper and lower ends of this segment are determined by the first sprouting of Δ as T_p grows, and therefore are easily computed. The result is that the waist has width $|A'B'|r/(p\delta+r)$. This also tends to 0 as p increases.

4. Further improvement. The upshot of Sections 2 and 3 is that without sacrificing either simple connectivity or the partial Kəkeya property, the set $K^{(0)}$ shown in Figure 2 can be replaced by a set $K^{(1)}$ like that shown in Figure 3 (using a thin tree of suitably large index) so that the area of $K^{(1)}$ is slightly smaller than that of $K^{(0)}$. More exactly, the small triangle Δ is still there, but the area outside has been reduced almost by the factor $r<1$. The next step is to reduce the area still more by doing the same construction again. The bulk of the area of $K^{(1)}$ resides in the joins $J_1, \cdots, J_m$. Let us subject each J_i to the same treatment as was given to J.

In fact, what has been said in Section 2 can be repeated verbatim with no more change than adding the subscript i to everything in sight, and replacing xy by the new coordinate system

$$x' = -x+\delta$$
$$y' = -y.$$

This puts J_i in the position shown for J in Figure 2. The points A_i' and B_i' corresponding to A' and B' in Figure 2 but not shown in Figure 3 are obtained by extending A_iC_i and B_iC_i to where they meet the y'-axis, which is $x=\delta$. They are in the waist of T so that the triangle $\Delta_i=A_i'B_i'C_i$ which corresponds to Δ in Figure 2 is contained in T.

Applying the construction of Sections 2–3 to J_i replaces J_i by a tree T_i extending from $x=\delta$ to $x=-r$ and a collection of new joins $J_{i1}, \cdots, J_{im_i}$ extending from $x=\delta$, to $x=1$. This is done for each $i=1, \cdots, m$, and our new partial Kakeya set is the union $K^{(2)}$ of the original tree T, m new trees T_i, and a doubly indexed family of new joins J_{ij} $(i=1, \cdots, m, j=1, \cdots, m_i)$. The motion of the needle from $A'A$ to $B'B$ in $K^{(2)}$ is the same as has been described for $K^{(1)}$, except that to get from $A_iC_iA_i'$ to $B_iC_iB_i'$ you no longer simply turn around C_i but rather go through an elaborate dance like the one described for $K^{(1)}$, using the elementary triangles of T_i and the joins J_{ij} alternately and in turn.

Area? As usual we write a_{ij} for the area of J_{ij}. Then $\sum a_{ij} = \sum ra_i = r^2a$. The total area of all the new trees T_i can be made arbitrarily small, and we conclude that the area of $K^{(2)} - \Delta$ can be made as near r^2a as we please.

The proof of Theorem 1 has two tricky places. The first was the tree construction which we owe to Besicovitch, as already described. The second is seeing that simple connectedness is preserved when the replacement of a triangle by a tree and joins is repeated as here. This is not obvious because (a) the joins J_{ij} for the new trees fall on top of an old tree T, already complicated, which they may criss-cross and (b) several triangles J_i are being treated simultaneously, and their respective replacements, trees and joins, 'may interact unpleasantly with each other. Postponing as before details of a proof, I give here somewhat informally the geometric reasons why $K^{(2)}$ is actually simply connected.

Note first that the second-generation trees $T_1, \cdots, T_m$ on the left do not interfere with each other. We have already seen that the triangles $J_1, \cdots, J_m$ they replace are disjoint. The vertices $C_1, \cdots, C_m$ of these triangles being distinct, by taking the orders of the trees large enough we can make their waists disjoint segments of $A'B'$. Then $T_1, \cdots, T_m$ will intersect each other only inside T.

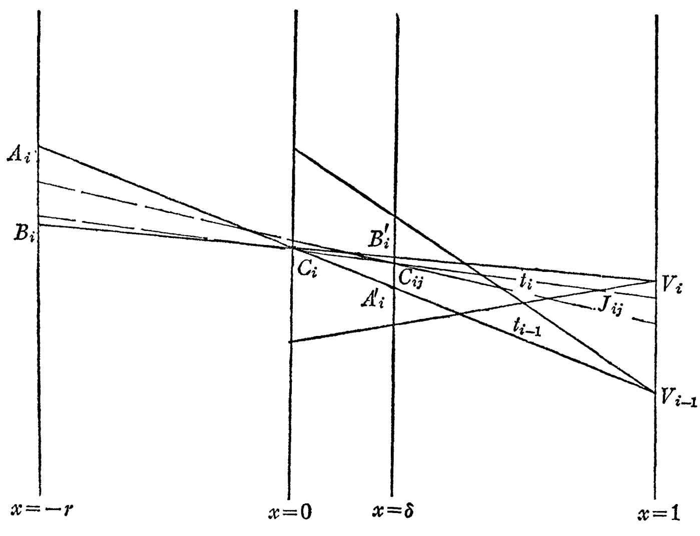


FIG. 6

We shall examine next the relationship between any second-generation join J_{ij} and the original tree T. In Figure 6 only part of T is shown, namely the union of the two elementary triangles t_i and t_{i-1} for the i in question, i.e., those which are joined by J_i. The base $A_i'B_i'$ of the triangle Δ_i from which T_i sprouts is contained in $t_i \cup t_{i-1}$, and therefore so is C_{ij}, which is between A_i' and B_i'. Comparison of slopes of lines in the figure then shows that J_{ij} is contained in the trape-

zoid $A_i'B_i'V_iV_{i-1}$. Now V_i and V_{i-1} are adjacent vertices on $x=1$ of the tree T. Thus J_{ij}, once it leaves T, never will meet T again. Moreover we already know from Section 3 that two joins J_{ij} and J_{ik} for the same i are disjoint, and from what we have just shown, two such joins for different first indices i, being contained in different trapezoids, can intersect each other only inside T. Putting these facts together, we see that the whole figure $K^{(2)}$ is "combed": The various tentacles (elementary triangles of T and second generation joins) do not cross each other outside of T so as to surround part of the complement of $K^{(2)}$.

What can be done twice can be done any number of times. We repeat the replacement of joins by trees and more joins as many times as necessary to achieve a preassigned minuteness of area. We proceed as follows.

Let $\epsilon>0$ be given, as well as the partial Kakeya set $K=K^{(0)}$ shown in Figure 2. We intend to make, by N repetitions of the replacement construction, a partial Kakeya set $K^{(N)}$ containing Δ such that the area of $K^{(N)}-\Delta$ is less than ϵ. To determine how large N needs to be, we budget $\frac{1}{2}\epsilon$ for all the Nth generation joins. We know that their total area will be r^Na, so we make $r^Na<\frac{1}{2}\epsilon$ by choice of N. (This is possible because $r<1$ and $\epsilon>0$.) We have left $\frac{1}{2}\epsilon$ for all the trees of N generations (excluding the area of Δ).

Now regardless of how many trees there are in the nth generation, their total area contribution is less than $2a^{(n)}/p$, where $a^{(n)}$ is the total area of the $(n-1)$st generation joins they replace, and p is their order. Since $a^{(n)}\leqq a$ for all n, by always taking $p>4Na/\epsilon$ we achieve the desired result.

That the needle can move from $A'A$ to $B'B$ in $K^{(N)}$ and that $K^{(N)}$ is simply connected require proofs by induction on N, the relevant basic considerations required being those already given for $N=2$.

5. Small Kakeya sets. Returning to Theorem 1 and Figure 1, we make a Kakeya set satisfying the conditions of Theorem 1 by applying the construction of the last three sections Q times, each time replacing one of the triangles J_q by a collection of trees and joins. The prickly part of each such replacement lies in the triangle J_q and the trapezoidal region opposite it, between J_{q-1} and J_{q+1}. Thus the union of all of them and the nucleus Π will be a simply connected set.

This set consists of Π and lots of triangles radiating out from Π, some of them elementary triangles of trees, and some of them joins. It would be interesting to have an estimate of the smallest number of such triangles required to make the area of the set less than the Bloom-Schoenberg number, which up to now is the record for simply connected Kakeya sets. This is not easy because of the various ways in which ϵ can be divided up. If the ϵ-utilization indicated in the proof (which is not necessarily the best) is followed, you arrive at the following approximate values: $\epsilon=.284$, $\delta=.495$, $Q=3$, $N=6$; total number of elementary triangles and joins, about 10^{43}.

The reader is invited to draw the figure!

6. Connectivity. For the construction in the preceding four sections to constitute a proof of Theorem 1, it remains to prove that the resulting Kakeya sets

are simply connected. The sets being as complicated as they are, the proof requires some care, and this part will be correspondingly technical. Even so, many details will of necessity be omitted, with indications of how they can be supplied. Without so stating in each lemma, I reserve the right to increase the order of the trees involved when that makes the proof easier, even though I believe all the lemmas are true as stated, without restrictions on p.

However complicated they may be, all the sets considered in the proof of Theorem 1 are polygons, so that in this section we need not be concerned with the subtleties of point-set topology. From now on the term "**set**" shall mean closed polygonal set in the plane, and "**path**" shall mean polygonal path.

A set is **connected** if and only if each pair of points in the set can be joined by a path in the set. It follows easily that if A and B are connected and $A \cap B$ is not empty, then $A \cup B$ is connected. It then follows by induction that, more generally, if $A_1, \cdots, A_n$ are connected, and if $A_i \cap A_j$ is not empty for all i, j, then $\bigcup_i A_i$ is connected.

A set is **simply connected** if and only if it is connected and each closed path in the set can shrink continuously to a point in the set. The technique for proving complicated sets simply connected is based on the following two facts analogous to those given in the last paragraph for connectedness.

LEMMA 1. *If A and B are simply connected, and if $A \cap B$ is connected and not empty, then $A \cup B$ is simply connected.*

Proof: This is the simplest special case of the Seifert-van Kampen theorem [5]. Since the general theorem is considerably more complicated, and I have no reference for the special case, I shall sketch the proof.

Let γ be any closed path in $A \cup B$ beginning and ending at some point P_0 in $A \cap B$. The first step is to find a finite sequence of points $P_1, \cdots, P_n$ in order on γ such that for each $i=1, \cdots, n$ the arc of γ from P_{i-1} to P_i (and also the last arc from P_n to P_0) is contained in either A or B. Then A and B being closed, the division points P_i are in $A \cap B$. Since $A \cap B$ is connected, each P_i can be joined to P_0 by a path β_i contained in $A \cap B$. This decomposes the original closed path γ into $n+1$ closed paths $\gamma_1, \cdots, \gamma_{n+1}$, where γ_i starts at P_0, goes to P_{i-1} along β_{i-1} (or along γ, if $i=1$), thence to P_i along γ, and thence back to P_0 along β_i (or along γ, if $i=n+1$). Each of these closed paths is contained in either A or B, and so can be shrunk to P_0, in A or in B, because A and B are simply connected. Doing this simultaneously for all γ_i shrinks γ to P_0 in $A \cup B$.

LEMMA 2. *Suppose* (i) *$A_1, \cdots, A_n$ are simply connected*; (ii) *$A_i \cap A_j$ is connected for all i, j*; (iii) *$A_i \cap A_j \cap A_k$ is not empty for all i, j, k; then $\bigcup_i A_i$ is simply connected.*

Proof: For $n=2$ this reduces to Lemma 1. Apply induction on n. The inductive hypothesis is that $B = \bigcup_{i=1}^{n-1} A_i$ is simply connected; by (i) A_n is also. Now

$$A_n \cap B = \bigcup_{i=1}^{n-1} (A_n \cap A_i)$$

is connected and not empty, because by (ii) $A_n \cap A_i$ is connected for all i and by (iii) the intersections of these sets in pairs are not empty. Thus $A_n \cup B = \bigcup_{i=1}^{n} A_i$ is simply connected by Lemma 1, completing the induction.

LEMMA 3. *Thin trees are simply connected.*

Proof: Induction on the order p of the tree. The lemma is trivial for $p=0$. Assume that T_{p-1} is simply connected. For each elementary triangle t_i of T_p let $A_i = T_{p-1} \cup t_i$. Then Lemma 2 applies to the union $\bigcup_i A_i = T_p$ because: (i) Lemma 1 applies to the union $T_{p-1} \cup t_i$ to show A_i is simply connected; (ii) $A_i \cap A_j = T_{p-1}$ for $i \neq j$, and (iii) T_{p-1} is contained in all A_i. (Geometric details left to the reader.) Lemma 2 then completes the induction. (We could prove by a different method that the more general trees in Section 1 are simply connected.)

The steps needed to prove our Kakeya set is simply connected are best brought out by a direct attack. The Kakeya set K is the union of Q sets K_q of the type $K^{(N)}$ of Section 3. To show that Lemma 2 is applicable to this union the main task is proving hypothesis (i), that each K_q is itself simply connected. Let us dispose of the other hypotheses first; (iii) is trivial, since every K_q contains the center of Π; (ii) is likewise trivial, except when the indices of the K_q's being intersected are consecutive. That $K_q \cap K_{q+1}$ is connected for each q can be seen by careful examination of a figure; the reader can supply a proof modeled on the proof of Lemma 6 below.

From now on we are concerned with showing that a single $K_q = K^{(N)}$ is simply connected. We use the coordinate system of Figure 2, and set $K^{(N)} = L \cup R$, where L is the part of $K^{(N)}$ in the left half-plane $x \leqq 0$, and R is the part in $x \geqq 0$. If we prove that L and R are simply connected, then Lemma 1 will apply to show that $K^{(N)}$ is simply connected. Now L decomposes naturally as the union of m sets (m being the number of first-stage joins), each consisting of the segment $A'B'$ on $x=0$ and the part of L descended from one first-stage join by all subsequent replacements. Each of these sets is structurally like R. If we prove R is simply connected, then the proof will apply to each of these as well, and Lemma 2 will apply to show that L is simply connected also.

From now on, then, we are concerned to show that the part of $K^{(N)}$ in $x \geqq 0$ is simply connected. The cases of odd and even N are not identical, and both must be treated, so that the result will apply to L as well as R. The techniques of proof are the same for the two cases, however, and we shall give only the proof for odd N, so that only trees are involved, the final joins being on the other side. The following notations are needed. $K^{(N)}$ is the union of trees from all the odd stages of replacement from 1 to N. Each tree of the kth generation ($k=3, 5, \cdots, N$) is recognized by its $k-1$ indices, the first $k-3$ of which are the indices of the tree of stage $k-2$ from which it is descended by two replacements. (In the first stage there is just one tree T, ancestor of them all.) If T' and T'' are two of these trees, we shall say that T'' is **descended** from T' when the index system of T'' begins with the index system of T'. For each Nth generation tree

T' there is a unique chain of $(N+1)/2$ trees, one from each odd-numbered generation, starting with T and ending with T', such that each tree in the chain is descended from all its predecessors. (The chain is obtained backwards from T' by peeling off indices two by two.) We shall call the union of all the trees of the chain $S(T')$. We intend to prove that $K^{(N)}$ is simply connected by applying Lemma 2 to the union $R = \bigcup S(T')$ (T' ranging over all trees of the Nth generation).

If t is any triangle with one vertical side, it is convenient to write $\Lambda(t)$ to mean the *slope-interval* of t, that is the set of real numbers which are slopes of lines joining the remaining vertex of t to points of the vertical side. In its tortuous motion from line $A'A$ to line $B'B$ in $K^{(1)}$ the needle has a slope which traverses in succession the intervals $\Lambda(t_0)$, $\Lambda(J_1)$, $\Lambda(t_1)$, $\cdots$, $\Lambda(J_m)$, $\Lambda(t_m)$. Thus these intervals in the order given constitute a partition of $\Lambda(\Delta)$. Refer again to Figure 3: with each J_i is associated another triangle $J_i' = V_{i-1}C_iV_i$ having the same slope interval. The part of J_i' not in T is a convex region U_i bounded on the right by the segment $V_{i-1}V_i$ and on the left by a polygonal arc from V_{i-1} to V_i which is part of the boundary of T, composed of segments of the sides of elementary triangles. Since the slope of this bounding arc is never in $\Lambda(J_i')$ $=\Lambda(J_i)$, any ray originating in U_i and extending to the right with a slope in $\Lambda(J_i')$ does not meet T, a fact which will be useful in proving Lemma 6. Let T^* be the union of T and the regions $U_1, \cdots, U_m$, and similarly for other trees. The following lemma repeats remarks made informally at the end of Section 4.

LEMMA 4. *$T_{ij}^* \subset T \cup U_i$ for all i, j.*

Proof: (See Figure 6.) The triangle Δ_i from which T_i grows has its vertex C^i strictly between A' and B' and its base $A_i'B_i'$ in the waist of T. (There is one index i, the middle one, for which $A_i'B_i'$ is all of the waist of T.) This triangle is contained in T, and if we take the order of T_i large enough, then the part of T_i in the strip $0 \leqq x \leqq \delta$ is also contained in T. Repeating these remarks at the second stage we conclude that the part of T_{ij} in the same strip is contained in T_i, and hence in T. Now any point in T_{ij}^* with $x > \delta$ is reached from the waist of T_{ij} by a ray to the right with slope in $\Lambda(\Delta_{ij})$, where Δ_{ij} is the triangle from which T_{ij} grew. The waist of T_{ij} is contained in the segment $A_i'B_i'$, which is contained in J_i'. Also

$$\Lambda(\Delta_{ij}) = \Lambda(J_{ij}) \subset \Lambda(J_i) = \Lambda(J_i').$$

Starting in J_i' and moving to the right on a ray with slope in $\Lambda(J_i')$, you stay in J_i' until $x = 1$. This proves that the part of T_{ij}^* in $x > \delta$ is contained in J_i', which in turn is contained in $T \cup U_i$, completing the proof.

COROLLARY 1. *If T' is a tree (of any odd generation) whose first index is i, then $T' \subset T \cup U_i$.*

Proof: Repeated application of Lemma 4.

COROLLARY 2. *If T' and T'' are trees whose first indices are different, then $T' \cap T'' \subset T$.*

Proof: If $i \neq j$, then U_i and U_j are disjoint. Apply Corollary 1.

LEMMA 5. *$S(T') \cap S(T'')$ is connected for every pair T', T'' of N-th stage trees.*

Proof: Let T''' be the "youngest common ancestor" of T' and T''. That means the indices of T''' are the common indices of T' and T'' from the beginning, up to the first place, where they disagree. Applying Corollary 2 of Lemma 4 to T''' we see that $T' \cap T'' \subset T'''$, in fact more: $S(T') \cap S(T'') \subset T'''$. Thus $S(T') \cap S(T'') = T'''$. This tree is certainly connected. (In fact simply connected, according to Lemma 3.)

LEMMA 6. *$T \cap T'$ is connected for every T'.*

Proof: Let the first index of T' be i, so that $T' \subset T \cup U_i$. We know that $T \cap T'$ contains the segment σ' of the y-axis which is the left boundary of T'. We shall show that each point P in $T \cap T'$ is joined to σ' by a segment in $T \cap T'$. Indeed, P is in some elementary triangle of T', and so is joined to some point Q of σ' by a segment whose slope λ belongs to the slope interval of that triangle, and hence to $\Lambda(J_i)$. If the segment PQ were not contained also in T, it would have a point in U_i, and then P would be on the ray extending to the right from that point with slope λ. This is impossible because $P \in T$.

COROLLARY. *If T'' is descended from T', then $T' \cap T''$ is connected.*

Proof: Apply Lemma 6 to T'.

LEMMA 7. *$S(T')$ is simply connected for every N-th stage tree T'.*

Proof: Let $T = T^1, T^2, \cdots, T^M = T'$, where $M = (N+1)/2$, be the chain for T', so that $S(T') = \bigcup_{k=1}^{M} T^k$. We apply Lemma 2 to this union. Each T^k is simply connected by Lemma 3; their intersections in pairs are connected by the Corollary of Lemma 6; and their intersection as a family is not empty because their bases on the y-axis form a nested sequence of segments.

We have now assembled all that is required to prove that R is simply connected, and so complete the proof of Theorem 1. In applying Lemma 2 to the union $R = \bigcup S(T')$, Lemma 7 gives hypothesis (i), Lemma 5 gives (ii), and (iii) is trivial, since $T \subset S(T')$ for all T'.

7. Star-shaped Kakeya Sets. Turning to the proof of Theorem 2, let K be a Kakeya set of area a which is star-shaped with respect to the origin O. Because K is star-shaped, K contains for each position of the needle in it the whole triangle Δ whose vertex is O and whose base is the needle. The area of Δ is $\frac{1}{2}\delta$, where δ is the perpendicular distance from O to the line of the needle, but we cannot immediately add the areas of such triangles to estimate the area of K because they will overlap. Let Γ be the fixed circle with center O and radius $1/6$, and let Δ' be the part of Δ outside Γ. The strategy of the proof is to select needle positions to give disjoint sets Δ' with as large an area as possible. We can then show that if the total of these areas outside Γ turns out to be too small, the area of the part of K inside Γ is necessarily large, so that combining the area outside and the area inside, we always get at least the value $\pi/108$ in the theorem.

Let the needle turn continuously around in K, the motion being parametrized by an independent variable t. At any instant t the needle lies on a line which is described by its direction λ and its distance δ from O, both continuous functions of t. (The line is not oriented, and λ is taken to be a radian angle modulo π.) It will be convenient to designate corresponding values of t, λ, and δ (also Δ and Δ') by affixing the same subscript to each.

The area of Δ' depends not only on δ, but also on the position of the needle along the line. For fixed δ, the needle position which minimizes this area is the one for which Δ is isosceles. The area of Δ' for this worst case is a known but complicated function of δ, which for small δ is approximately $2\delta/9$. To get precisely the lower bound stated in the theorem, we need the following sharper estimate.

LEMMA. *The area of* Δ' *is at least* $\frac{1}{27}\sin^{-1} 6\delta$.

The proof of the lemma is a lengthy, ugly, but elementary computation (which I omit) starting from the exact expression for the area. Only values of δ in the interval $[0, \pi/54]$ need be considered, because $\delta > \pi/54$ would immediately give a Δ of area $\frac{1}{2}\delta > \pi/108$.

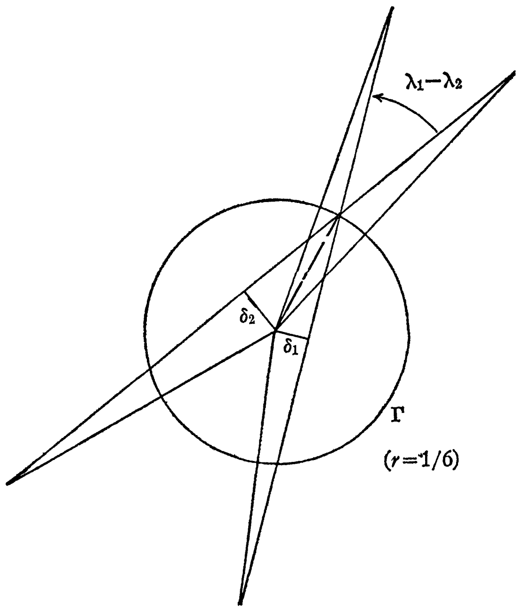


FIG. 7

If at two times t_1 and t_2 the directions λ_1 and λ_2 are different enough, then Δ_1' and Δ_2' will be disjoint. From Figure 7 it will be seen that the sufficient condition for this is (mod π) $|\lambda_1-\lambda_2| \geqq \sin^{-1} 6\delta_1+\sin^{-1} 6\delta_2$. We proceed to select a sequence of times $t_1, t_2, \cdots$ such that $\Delta_1', \Delta_2', \cdots$, are disjoint. First choose t_1 so as to give δ its maximum value δ_1. Let I_1 be the open interval of directions

λ centered at λ_1 and $4 \sin^{-1} 6\delta_1$ wide. Then for all t such that $\lambda \notin I_1$, since $\delta \leqq \delta_1$, we have $|\lambda - \lambda_1| \geqq 2 \sin^{-1} 6\delta_1 \geqq \sin^{-1} 6\delta + \sin^{-1} 6\delta_1$, and therefore Δ' and Δ_1' are disjoint. Next select t_2 from the set of times (which is closed) where $\lambda \notin I_1$ so as to maximize δ_2. Let I_2 be the open λ-interval centered at λ_2 and $4 \sin^{-1} 6\delta_2$ wide. Select t_3 from the set of times where $\lambda \notin I_1 \cup I_2$ so as to maximize δ_3, and so on, continuing thus as long as possible.

Suppose first that for some N the intervals $I_1, \cdots, I_N$ cover the whole λ-interval $[0, \pi)$ so that selection of a t_{N+1} is not possible. In that case the sum of the lengths of $I_1, \cdots, I_N$ must be at least π. Then the area of K is at least the sum of the areas of the disjoint sets $\Delta_1', \cdots, \Delta_N'$ contained in it; hence by the lemma

$$a \geqq \sum_{n=1}^{N} (1/27) \sin^{-1} 6\delta_n = (1/108) \sum_{n=1}^{N} 4 \sin^{-1} 6\delta_n \geqq \pi/108,$$

and Theorem 2 holds for this case.

The remaining case is when no finite set of the intervals $I_1, I_2, \cdots$ covers $[0, \pi)$, and the selection goes on forever. In this case we get an infinite sequence $\Delta_1', \Delta_2', \cdots$ of disjoint subsets of K and the lemma implies that the infinite series $\sum_{n=1}^{\infty} (1/27) \sin^{-1} 6\delta_n$ converges to a sum $b \leqq a$. Hence $\delta_n \to 0$. The union of all the intervals $I_1, I_2, \cdots$ is an open subset of $[0, \pi)$ of measure at most $\sum_{n=1}^{\infty} 4 \sin^{-1} 6\delta_n = 108b$. Its complement C in $[0, \pi)$ is then a set of measure at least $\pi - 108b$. For each t such that $\lambda \in C$, t was eligible for selection at every stage, and since δ_n was always maximum, $\delta \leqq \delta_n$ for every n. Since $\delta_n \to 0$ it follows that $\delta = 0$ for all such t. Now K being a Kakeya set, every λ is the direction of the needle at some time. Thus K contains a unit segment through O in every direction of the set C. The union of these segments has inside Γ area $\frac{1}{2}(1/6)^2$ times the measure of C, that is, at least $(1/72)(\pi - 108b)$. Adding to this our estimate for the area outside Γ, we obtain $a \geqq (1/72)(\pi - 108b) + b = \pi/72 - \frac{1}{2}b \geqq \pi/72 - \frac{1}{2}a$. Solving this inequality for a leads to $a \geqq \pi/108$, completing the proof.

REMARK: The continuity of the motion of the needle is not used in an essential way. With only small changes the same proof gives the following more general result: any star-shaped set which contains unit segments in all directions has plane outer measure at least $\pi/108$.

References

1. H. J. van Alphen, Uitbreiding van een stelling van Besicovitsch, Mathematica (Zutphen) B, 10 (1942) 144–157.

2. A. S. Besicovitch, On Kakeya's problem and a similar one, Math. Z., 27 (1928) 312–320.

3. ———, The Kakeya Problem, this MONTHLY, 70 (1963) 697–706.

4. F. Cunningham, Jr. and I. J. Schoenberg, On the Kakeya Constant, Canad. J. Math., 17 (1965) 946–956.

5. W. S. Massey, Algebraic Topology: An Introduction, Harcourt, Brace and World, 1967.

6. J. Pál, Ein Minimumproblem für Ovale, Math. Ann., 83 (1921) 311–319.

7. R. J. Walker, Addendum to Mr. Greenwood's paper, Pi Mu Epsilon J. (November 1952), 275.

PUBLIC UNDERSTANDING OF SCIENCE AND ITS IMPLICATION FOR MATHEMATICS

L. J. PAIGE, University of California, Los Angeles

I wish to thank Professor Burgess and the Committee on Arrangements for the opportunity to address you today. It's been a long wait; but, having helped to build this auditorium, literally, having graduated from this stage, I have finally made it to the rostrum. In addition to Professor Burgess' comments on the tasks which I have undertaken in the interests of mathematicians and mathematics, I think it only right that I should comment briefly on my other concerns.

I, as Dean of the Physical Sciences at UCLA, have had my problems this year. These have included a penurious legislature that insisted on the return of faculty positions and fiscal support for a discontinued summer quarter. And there are few administrators so fortunate as to have had three ROTC units under their immediate supervision during the past two years.

My burdens as Chairman of the Los Angeles Division of the Academic Senate have included a well-known Acting Assistant Professor; a running engagement with student activists and their concern for participation in University governance; and a legislature that threatened to remove all funds for the operation of the Senate from the budget. Perhaps I will be known simultaneously in years to come as that mean penny-pinching Dean and the last of the big time spenders.

In view of the preceding, it is difficult for me to pretend that I speak to you as a mathematician when, in fact, I feel that I return here today as "a defrocked monk at a meeting of the college of cardinals."

It would serve no useful purpose for me to define precisely the science I shall refer to when I speak of "public understanding of science." One of the difficulties for an understanding of science is that of language; the distinctions between basic scientific research, applications and technology are occasionally blurred for us specialists and surely most, if not all, of the time for the public. However, as a matter of convenience, my remarks will be restricted to the life, mathematical and physical sciences together with their interdisciplinary offspring; for example, biochemistry, geophysics, etc.

After two decades of spectacular growth of federal support for both basic and applied research, the funding of science is being seriously questioned by various segments of society. The recommendations and guidelines for science policy and federal support of academic science set forth by Vannevar Bush in 1945 in *Science: The Endless Frontier* are no longer sacrosanct. The National Academy of Sciences, sensing the seriousness of the situation, recently conducted a Symposium on "The Crises in Federal Funding of Science."

The problems are serious, not hopeless, and I refuse to be the present day version of Henny-Penny. First, let me note some reasons for optimism. The National Science Board in January stated the basic tenets of United States science policy to be [11, p. ix]:

"a. The United States will strive to remain competitive at or near the forefront of each of the major areas of science and, to this end, will continue to identify and support scientific excellence.

b. The Nation is committed to the principle that every young person should have the opportunity to pursue education to the extent of his ability and motivation irrespective of geographic origin or economic means.

c. The Federal Government has a responsibility to ensure that new scientific knowledge is utilized as rapidly and efficiently as possible in support of national goals and for the welfare of the world's peoples."

These basic tenets are supported, in essence, in the April 1970 report of the President's Task Force on Science Policy entitled *Science and Technology: Tools for Progress.*

Let us not deceive ourselves into believing that the recommendations of these reports will still the cries for social responsibility from our colleagues; satisfy the demands for relevance from our students; or, thwart the search for budgetary savings by our politicians. We are confronted with an educational problem of considerable magnitude.

I believe that the present unfavorable, indeed hostile, climate for science has been brought on by our own careless disregard of early warning signals. Hence, I shall occasionally point out past omissions, while I discuss briefly the understanding of science, or apparent attitudes, of several segments of our society; the layman, our nonscientist colleagues, the students and politicians. After that, the implications for mathematics will be seen in the proper context and we can visualize the steps that should be taken to preserve the essential funding.

1. The Layman's Understanding of Science. It is my conviction that only in the last few decades has the layman been called upon to understand science or the role played by science in technology and society.

In the past, the intellectual curiosity of men like Galileo and Kepler required only the enlightened support of a few who were convinced that an expanding knowledge of nature would contribute to the well-being of mankind. The contributions of Newton, Pasteur, Faraday and others to scientific progress during the industrial revolution were largely efforts supported by academic or government positions. The layman was assured that the discoveries of science would lead to practical applications and industrial progress with concomitant improvements in his welfare. Did he really understand that it is only in the increasingly sophisticated technology of the last century that basic scientific research has played a significant role for industry?

Where does the layman's understanding of science come from? If we assume that a great deal of understanding derives from popular scientific journals, then our efforts have been diluted often with misrepresentation at best and with technological hucksterism at worst.

A classic example of the latter, noted with the aid of twenty-twenty hindsight, occurs in Dubos' "Reason Awake" where he quotes from the July 1899 issue of *Scientific American* [**2**, p. 95]:

"The improvement in city conditions by the general adoption of the motor car can hardly be overestimated. Streets clean, dustless and odorless, with light rubber-tired vehicles moving swiftly and noiselessly over their smooth expanse, would eliminate a greater part of the nervousness, distraction and strain of modern metropolitan life."

I would like to assure you that equally preposterous statements can be found in the literature of the past few years with respect to computers. The waiting of days and days by voters in several of our California cities to learn the results of recent elections provides little assurance to the layman that science and technology are cooperating for his benefit.

In a more serious vein, even if the layman is adequately prepared to read the technical articles of scientific journals, the probability is that his interests will be highly selective except for the editorial comment, letters, and the news articles concerned with science policy. In these, more often than not, the public is treated to excesses in the scientific community. Remember MOHOLE or the controversy surrounding the selection of WESTON, ILLINOIS as the site for the NATIONAL ACCELERATOR LABORATORY?

If we think we are engendering the layman's support and understanding for science in our discussions in journals, it is appropriate to note what President Philip Handler of the National Academy of Sciences had to say in a recent speech [12]:

"Meanwhile, how can we effectively protest the current injury to the national research endeavor, when, inevitably, that must seem to be defense of our own personal incomes be they summer salaries, partial or complete annual salaries. For an illustration of such arguments at their hypocritical worst, I refer you to a letter in *Science* for March 27, from a group of mathematicians. Inevitably our distress signals must fall on deaf ears when our arguments can quite accurately be described as personally self-serving, regardless of the honesty or objectivity of our intentions."

Perhaps, and it might be just as well, the scientists are talking to each other in the journals and the layman turns elsewhere for his information. Where?

Science Fiction? The prospect of Orwell's "1984," Bradbury's "Fahrenheit 451," Huxley's "Brave New World," or Pohl's "The Case Against Tomorrow" being the basis for an understanding of science is hardly encouraging. Moreover, the analysis of the latest science fiction in Mark R. Hillegas' "The Future as Nightmare: H. G. Wells and the Anti-Utopians" emphasizes the change in viewpoint from that found in H. Bruce Franklin's "Future Perfect: American Science Fiction of the 19th Century."

Popular Nonfiction? It is difficult for me to judge the effectiveness of the many popularizations of scientific work; the efforts of *Life* magazine in their Science Library series; or the reception of books like Don K. Price's "The Scientific Estate" and Daniel S. Greenberg's "The Politics of Pure Science." However, when it comes to alerting the public to problems of our environment, there is little doubt that Rachel Carson's "Silent Spring" had a profound effect.

Television? I don't believe that the many medical shows with their displays of pseudo-operations and the sophisticated monitoring equipment, or the ingenious technological devices of MISSION IMPOSSIBLE contribute to an understanding of science. And the viewers of the early morning classrooms of science will, at best, provide a basis for support in the distant future. Fortunately, the public concern for our space efforts, following the courageous (albeit questionable) goal setting by our former President, has permitted science and technology to display an incredibly sophisticated skill; but, I wonder how many laymen now question the cost of our scientific advances in view of the other problems facing society?

To conclude, I can only infer that the layman develops attitudes about science and not an understanding. These attitudes are subject to shock and it is not necessarily true that the resulting changes are always favorable.

The moral, intellectual and scientific reverberations of Alamagordo still abound. The consequences of SPUTNIK were indeed a decade of magnificent abundance for science and science education. On the other hand, Rachel Carson's "Silent Spring" and Ralph Nader's "Unsafe at Any Speed" revealed glaring faults to the layman in the halo of science and technology.

Today, student disorders and faculty dissension are providing another shock not understood by laymen. Higher education, not merely academic science, will feel the strain.

Before I turn to the task of suggesting steps which may be taken to refocus the picture of science which we ourselves have permitted, let's first analyze the attitudes towards science by other crucial segments of our society.

2. Our Academic Colleagues and Science. There is no particular need for me to review and analyze the communication gap which is said to exist between scientists and non-scientists in the academic community. C. P. Snow's famous lecture of 1959, "Two Cultures and the Scientific Revolution," seems to have attracted enough attention. The antagonists from both camps still flail away, as can be noted in the TIMES LITERARY SUPPLEMENT of April 23 and 30, 1970 (F. R. Leavis and Noël Annan). It is my contention that the widening gap of misunderstanding has been overemphasized and, by now, this controversy could well indulge in a heavy dose of Moynihan's "benign neglect."

The understanding of science by our nonscientific colleagues is an entirely different matter from that of the layman. They certainly recognize the excitement of discovery; the pursuit of knowledge for knowledge's sake without conscious concern for immediate application; and the desire to investigate problems of one's own choosing. However, their attitudes are tinged with cynicism when they note what Dubos writes [2, p. 234]: "Individual scientists exhibit intellectual integrity in their professional work, but as a group we tend to abandon intellectual and ethical discipline when unwarranted claims give promise of increasing social support for the problems we happen to find interesting. Dedication to the discovery of truth for truth's sake seems to be quite compatible

with the more mundane desire to work in fashionable fields, preferably those which are well financed and likely to be rewarded by academic promotion and glamorous prizes—let alone plump consultant fees."

Our nonscientist colleagues would be inhuman if they were not dismayed and jealous of the financial support afforded science since World War II. Moreover, the scientist's callous disregard for humanistic support, while we wallowed in research institutes, innumerable symposia (invariably at attractive places to visit), graduate fellowships, supported sabbatical leaves and lower teaching loads, was not designed to engender their understanding or concern for our present plight.

It is fortunate that President Philip Handler of the National Academy of Sciences is concerned over these imbalances and, in his discussion of a new Federal agency for research and higher education, argues as follows [**12**]:

"Some of the science community are already uncomfortable with the support of the social sciences provided by the NIH and the NSF. Undoubtedly, this group will be further dismayed by the proposal that a common agency also engage in the funding of scholarship in the humanities. Yet, because the humanities and the arts are coequal with science in our national culture, I consider it appropriate that the modest Federal programs in these areas be folded into the same agency. Collectively, these disciplines represent the intellectual quality of our national life and their success today will be crucial to our national viability tomorrow. What is true about imbalances within the sciences under the present system for science, where priority definition is so difficult, is even more true within learning as a whole where there is a great imbalance in the relative positions of the natural sciences, the social sciences, and the humanities. Although government support for scientific research is not what it should be, or what I believe the national interest to require, it is gigantic in comparison with government support for scholarship and education in the humanities. Educational and research dollars need not be distributed among the three major divisions of learning on an equal basis, for their needs are different. But the imbalance should not be so pronounced that the advancement of one area is attended by diminution in the attractiveness and morale of the other. Moreover, the present separation of government support for the humanities, social sciences, and natural sciences tends to institutionalize an artificial division that, in itself, is harmful to the national interest. One of the greatest present needs is to bring together the scientific and humanistic enterprises so that scientific discovery in the future will take place within the context of humanistic thought about how best to use the discoveries that are being made headlong in the sciences."

Let us hope that our nonscientist colleagues will be equal in their understanding of scientific research.

3. The College Student and Science. Anyone who is bold enough to attempt to predict or understand the attitudes of students today runs the risk of being wrong tomorrow. Nonetheless, I am willing to try.

First, there is no doubt that there is a growing alienation and disenchantment

of students with science and technology. This should come as no surprise to many science departments since the evidence was available long before the pragmatic effect of today's employment opportunities haunted doctoral candidates. For many years student enrollment in science classes, particularly those designed for majors in the physical sciences, has not kept pace with the increase in college population. Mathematics classes are an exception to the preceding observation; but the increasing emphasis on mathematics for the social and life sciences has contributed significantly to these favorable enrollment trends.

The students have directed their severest criticism of emphasizing research over teaching in promotion policies at the science departments. We don't have to wonder why, when many professors are seen only by graduate students. Similarly, the student protests directed at war-related research have undermined our assurances of the value of science for society. Now we can add the growing stories of technological waste and destructiveness as a target for student concern and, subsequently, further alienation from science.

I believe that a considerable portion of the student demand for changes in grading procedures and the abandonment of breadth requirements stems from their disenchantment with science and the supposed "irrelevance" of our courses. The questions of racial unrest, educational opportunities, deteriorating cities, pollution of rivers, dying lakes, smog-filled lungs, inadequate health care and a multitude of other social ills seem to students to be far removed from the absorbing intellectual demands of "elementary particles," "chemical isomers" or "Lebesgue integration."

I think the students are wrong. While they, and many laymen, attribute the deterioration of our environment to science (without distinguishing between science and technology), they fail to understand that solutions will *not* be found without significant contributions from science in both basic and applied research. It seems to me that quite the opposite from the point of view of students is true; namely, science education, as presently constituted, is notably inadequate to provide an understanding of the scientific and technological problems which must be considered in resolving many of our difficulties. Moreover, the establishment of priorities in our society requires an understanding of the contributions which science and technology can provide.

It seems to me that an adequate understanding of science is an indispensable necessity to any rational discussion of the moral questions raised by organ transplants, genetic manipulation, or environmental management. Otherwise, students may well formulate their attitudes on the basis of the most inappropriate rhetoric.

Let me give some simple illustrations. The physiological, psychological and social implications of "the pill" are all important aspects of population control; yet I dare say, the students are concerned almost exclusively with the latter. Nuclear energy is viewed primarily as a ghostly spectre that threatens civilization in the form of MIRV's and ICBM's: but it seems reasonable to assume that controlled thermonuclear fusion would be of the greatest benefit to man-

kind. Isn't some understanding of the research involved toward this end a necessity when priorities are being established for desirable social goals?

4. Politicians and Science. The support of basic and applied research in science, as well as of programs designed to strengthen scientific training throughout our educational system, was normally expected of our elected representatives in Congress during the two decades following World War II. Unfortunately, the last few years have witnessed a steady erosion of congressional goodwill towards science.

The original funding of science reflected the persuasive arguments of Dr. Vannevar Bush and his colleagues; but it is difficult to imagine that all members of Congress offered their support because of any substantial understanding of the social benefits to be derived. The contributions of scientists to the war effort were impressive. It is quite likely that most Congressmen were more persuaded by arguments stressing our need for scientific and technological strength for reasons of national security than moved by a conviction that basic scientific research is a fundamental need of society.

Later, the science advisors to successive Presidents were articulate spokesmen for an expanding support of both basic and applied research. Science found extremely capable and understanding men in Congress to advocate the value of scientific research to society. Today, many of these congressional supporters are gone and we might wish for more to express their position as positively as did Senator E. M. Kennedy in addressing his colleagues during hearings on the National Science Foundation [9, p. 64]:

> "For my part, I believe that the Administration is making a woeful error in cutting back substantially on basic scientific research. I recognize the need for budgetary belt-tightening in a period of inflation. But even in a period of inflation, a responsible government must choose its priorities wisely; it cannot slash every Federal program. And it is my belief that no government with a proper set of priorities can, at this moment in history, seriously consider reducing its commitment to scientific research. We are just on the verge of reordering our national priorities, of freeing more scientists to work on civilian rather than military problems. I do not understand how this can be the time to cut back on support for science and scientists."

I hope that Senator Kennedy will be able to persuade his colleagues. Have any of you written Senator Kennedy to add your reasoned comments to those who have testified before his subcommittee? You should.

Quite understandably, Congressional interests now turn to popular environmental issues, and the pleas for science support are muffled by the rhetoric of those distrustful of science and technology; a rhetoric colorful in its predictions of disaster while devoid of reasoned analysis. I do not mean to imply that the issues of ecology are unimportant but it must be emphasized to Congress that solutions to the problems raised will require significant research contributions from science.

The Mansfield amendment restricting support of research by defense agencies to mission oriented needs has influenced other funding agencies to emphasize applied research. This is a classic example of legislative fallout far beyond the intended target and, in this particular instance, quite incompatible with the recommendations of any committee concerned with the future of science. We can only wonder where politicians obtain their advice.

At a time when Congress and State Legislatures are being forced to establish priorities between the funding of education and other programs of social concern, we expect our professional societies to provide the rationale for research as an integral part of teaching at both the undergraduate and graduate level. Moreover, in this period of competing demands upon our resources, it behooves science to present forcibly and clearly the consequences of deferring support for basic scientific research.

The President's National Goals Research Staff in its report entitled "Towards Balanced Growth: Quantity with Quality" emphasized key areas in which there are existing debates over a growth policy: population growth and distribution, environment, education, basic natural science, technology assessment and consumerism. With respect to basic natural science, the summary states:

> "One of the major decisions with which we are faced is that of the level of support we will furnish basic science in the future. This is clouded by the problem of making basic research "useful" in the short run. . . . Setting research priorities on the grounds of probable utility is often a choice of possible short-term benefits against the longer-term ones which might result from a more rapid expansion of the basic pool of knowledge by permitting science to pursue the internal logic of its own development.
>
> What is needed, and may in fact be developing, is a forum in which the partially conflicting needs for maintaining the integrity of the core of basic research and the practical needs of society are resolved.
>
> In conjunction with the need to work out an appropriate level and distribution of funding, we must face the fact that an articulate minority are attacking the very rationale and spirit of science and of rational inquiry itself—the most elementary tools man has for the orderly guidance of his affairs."

The National Academy of Sciences—through its committees and the National Science Board will continue to exert considerable influence in the ensuing discussions. The President's Task Force on Science Policy has submitted its recommendations and, I would assume, our professional societies are reviewing their own needs in the light of these reports.

It is now time for you to bring to the attention of your elected representatives your views on the recommendations which have been made. It has been my experience this Spring that our elected officials do respond to individual expressions of concern and I would urge you to write the appropriate congressional committees.

5. The Implications for Mathematics. The growing concern of the public, our colleagues, students, and politicians that basic research in the sciences does not reflect adequate consideration for the social implications of discoveries, nor a direction of planning compatible with an increasing articulation of desirable social goals, will be science's burden for years ahead. Some have referred to the present questioning of research support as a crisis in the "social support system of science."

Let there be no doubt that the present crunch will continue. It takes little imagination to formulate a series of social ills in our society demanding immediate attention and presenting competing claims upon the nation's resources. Professor G. Birkhoff has put it quite bluntly [**10**], "The picnic is over."

Let us look at some of the consequences which the mathematical community must face. First, it is to be expected that a major portion of any additional funds recommended for the National Science Foundation this year will be assigned to interdisciplinary programs directed at the problems of society. Even a casual reading of the testimony before the Special Subcommittee on NSF of the Senate reveals this fact. Hence, the fiscal support available for fundamental research in science, including the mathematical sciences, will remain approximately the same as last year.

We must recognize the fact that the scientific community will be engaged in the assignment of priorities for funding among various fields and that the mathematics community will be called upon to prescribe its own internal priorities. With respect to the former assignment of priorities, we should expect the various reports prepared for the Committee on Science and Public Policy of the National Academy of Sciences to be the basis for the appeals to be made by each discipline. The reports on the mathematical sciences prepared by COSRIMS should serve us well; but, there will be other pressures.

The most widely discussed criteria proposed for the assignment of priorities to scientific research are those advanced by Dr. A. Weinberg of the Oak Ridge National Laboratories [**5**, p. 75]. The criteria of justification proposed for the support of science were: technological merit, scientific merit, and social merit. It is in the discussion of scientific merit that he states,

> "I would therefore sharpen the criterion of scientific merit by proposing that, other things being equal, *that field has the most scientific merit which contributes most heavily to and illuminates most brightly its neighboring scientific disciplines.*"

If the preceding is taken without modification as a reasonable basis for the allocation of funds within the National Science Foundation, then the mathematical sciences section will need all of the assistance the professional societies can provide for the justification of their requests. To illustrate my concern, I note that in Weinberg's discussion of scientific merit preceding his recommendation he appeals to the following comment of von Neumann:

> "As a mathematical discipline travels far from its empirical source, or still

more, if it is a second or third generation only indirectly inspired by ideas coming from reality, it is beset with grave dangers. It becomes more and more pure aestheticizing, more and more purely *l'art pour l'art*. This need not be bad if the field is surrounded by correlated subjects which still have closer empirical connections or if the discipline is under the influence of men with an exceptionally well developed taste. But there is grave danger that the subject will develop along the line of least resistance, that the stream, so far from its source, will separate into a multitude of insignificant branches, and that the discipline will become a disorganized mass of details and complexities."

To be brief, the appeal is to relevance; not in the sense attached to relevance by students but in the intellectual context of unifying concepts. I do not interpret von Neumann's remarks to be a clarion call for slavish devotion to the applications of mathematics; but I am uncertain that Weinberg and other scientists are not so inclined.

Many mathematicians have expressed the need for our courses and research efforts to reflect the relation between various areas of mathematics as well as to applications to other disciplines.

I would propose that our writing include more than a feeble pass at articles designed to illustrate the unifying aspects of abstract concepts for the nonmathematical scientist. The initial effort of the COSRIMS reports must be continued if we are to convince our scientific colleagues that the plea in Hardy's toast, "Here's to pure mathematics. May it never have any use," has not been fulfilled. Thus, I find articles of the nature of Saunders MacLane's in the June-July issue of the MONTHLY to be of considerable importance.

The extent to which we shall be able to influence the support of the mathematical sciences will depend, in part, upon the understanding, respect and support we provide our internal constituencies; applied mathematics, computer sciences, curriculum development, pure mathematics, and teacher education. It would be most unfortunate if we were to persist in internal bickering and end up seeking alms from our fiscally ravenous experimental colleagues.

I have devoted considerable time to what might appear to be the selfish interests of faculty members. Now I wish to consider the important component of our concern: the students. What will be the effect of present attitudes upon our students?

There is no doubt in my mind that the growing contention that science and technology are insensitive to our social problems is driving undergraduates from Science and Mathematics to the Social Sciences. This can only result in further alienation from mathematics and I submit that one of our curriculum disaster areas is in courses designed for nonmathematics majors in addition to those service courses we provide for the various disciplines.

Even if we choose to ignore the nonmathematics majors, our undergraduate majors cannot help but notice the reduction in fellowships and research assistantships for graduate study. It is estimated that the reduction will be approximately

20% this year. Is it any wonder that students are discouraged when the prospects for support are diminishing? And to this distressing note, we might add the publicity of an oversupply of Ph.D.'s which has been widely discussed in the mathematical community.

I believe that we have two problems regarding graduate students which demand our immediate attention. We must let Congress know that the reduction of graduate student support will seriously hamper the ability of the mathematical sciences to maintain the quality of research ability so necessary for the long range needs of science and society. The reasoned comments of the governing councils of our professional societies must be submitted to the appropriate congressional committees. It is not sufficient that we direct their attention to the COSRIMS reports.

I was somewhat surprised to read that Philip Handler has testified before a congressional committee that *all* graduate students, regardless of field of study, should receive government support. Frankly, since the present admissions policies for graduate study at many schools and departments do not reflect the problems of growth or the needs of society, I would be opposed to such a commitment at this time. There is no reason why the graduate schools of our nation should become an unplanned elitist counterpart of the depression's Civilian Conservation Corps. On the other hand, I do share Handler's desire to provide support for graduate study in the Arts, Humanities and Social Sciences in order to alleviate the incredible imbalance that has favored science for several decades. Our pleas should not be entirely self-serving since the future advances in science will have little meaning if the quality of scholarship in those areas is not recognized as important to our culture.

Our graduate students should also expect the mathematical community to continue to address itself to the character of our graduate programs. The discussions did not begin with COSRIMS or CUPM nor should they end with the articles now appearing by Professors Anderson, Herstein, Rosenberg and others.

Are we being honest with our graduate students if we continue to provide them with training and attitudes in a manner that society finds ill-suited for its needs? We can no longer ignore the priorities which society will impose upon science and mathematics. Our advanced training must acknowledge the probable type of employment of a majority of our doctoral candidates.

Finally, I come to the area of post-doctoral study and the implications I see for the mathematical sciences. The COSRIMS reports have recommended various minimal programs for the mathematical sciences and the National Academy of Sciences has issued a detailed analysis of post-doctoral education under the title "The Invisible University: Post-doctoral Education in the United States."

The governing boards of our universities have been largely unaware of the problems of postdoctoral education; postdoctoral support has been almost exclusively federal and foundational in character and the cost to the institution in space, equipment and faculty effort has been largely ignored. This is no longer the case.

Postdoctoral positions are seldom included in university budget requests and legislatures are questioning the need for space and equipment for people regarded by them to be temporary non-students. Moreover, the demands of undergraduate students for more teaching and less research emphasis have attracted legislative attention. We can expect embattled Appropriation Committees to increase their demands that faculty members devote more time in the classroom and less time to the research laboratories. Postdoctoral education could suffer.

The mathematical sciences must address themselves to articulating the need for postdoctoral education and the type of training most desirable. A particularly important aspect will be the distribution of available positions. It should be pointed out that almost 200 universities have postdoctoral positions but over one-half are at seventeen schools. In the present funding crisis, it is imperative that the postdoctoral opportunities of our major research centers not suffer by virtue of myopic demands by Congress for geographical distribution. It should be self-evident that the future of the mathematical sciences lies in the hands of our most distinguished scholars and the enthusiasm for research, both pure and applied, which they instill in our most promising young colleagues during their postdoctoral years.

There are other implications for the mathematical sciences which I could develop but I believe that enough have been presented to keep us busy. In closing, and not as an apology, I wish to point out that I have not emphasized the needs of our minority groups since their aspirations should receive attention in *all* of the persuasive dialogues I have suggested the mathematical sciences undertake. The ethnic free character of mathematics must not be used as an excuse for our failing to provide aid, opportunity, challenge and hope.

Presented on August 25, 1970 at the 51st Summer Meeting of the MAA in Laramie, Wyoming.

References: Books

1. Bush, Vannevar. 1960. Science; The Endless Frontier. National Science Foundation, Washington, D. C.

2. Dubos, René. 1970. Reason Awake: Science for Man. Columbia University Press, New York and London.

3. Greenberg, Daniel S. 1968. The Politics of Pure Science. New American Library, New York.

4. Price, Don. 1965. The Scientific Estate. Harvard University Press, Cambridge, Massachusetts.

5. Weinberg, A. 1967. Reflections on Big Science. The M.I.T. Press, Cambridge, Massachusetts.

Articles and Reports

6. MacLane, Saunders: Hamiltonian Mechanics and Geometry, this MONTHLY, Vol. 77 (1970) 570–586.

7. Mathematical Sciences: A Report. A report prepared by the Committee on Support of Research in the Mathematical Sciences (COSRIMS) for the Committee on Science and Public Policy. National Academy of Sciences Publication 1681 (1968).

8. Mathematical Sciences: Undergraduate Education. A COSRIMS report. National Academy of Sciences Publication 1682 (1968).

9. National Science Foundation Authorization, 1970. Hearings before the Special Committee on National Science Foundation, April 16 and 23, 1970. Government Printing Office.

10. Newsletter, Conference Board of the Mathematical Sciences, Vol. 5, No. 1, May 1970 (p. 15).

11. The Physical Sciences. A report of the National Science Board submitted to Congress 1970. Government Printing Office.

12. Science and the Federal Government: Which Way to Go. A speech delivered by Philip Handler, President, National Academy of Sciences to Federation of American Societies for Experimental Biology, April 14, 1970.

13. Science and Technology: Tools for Progress. A report on the President's Task Force on Science Policy. April 1970. Government Printing Office.

14. Toward Balanced Growth: Quantity with Quality. A report of the National Goals Research Staff to the President. July 4, 1970. Government Printing Office.

15. The Invisible University: Postdoctoral Education in the United States. A study conducted under the auspices of the National Research Council 1969. National Academy of Sciences, Washington, D. C.

MARKOV RANDOM FIELDS AND GIBBS ENSEMBLES

FRANK SPITZER, Cornell University

1. Introduction. There are two very interesting and apparently quite different ways of defining random configurations of points on a lattice (or so called random fields in the terminology of R. L. Dobrushin [1]). One of these is based on the formulation of statistical mechanics according to J. W. Gibbs. It is generally accepted as the simplest useful mathematical model of a discrete or lattice gas. Its physical significance was enhanced when it was shown to exhibit, in dimension $\nu \geqq 2$, the singularities associated with the phenomenon of phase transition. (See [2] and [3] for recent, mathematically rigorous treatments.)

The second class of random fields we shall consider is that of Markov random fields, introduced by Dobrushin [1]. It has no apparent connection with physics, being based instead on the most natural way of extending the notion of a Markov process with one dimensional, integer valued, time to the case of higher dimensional, lattice valued, time parameter.

The purpose of this article is to show that *these two ways of defining a random field are equivalent.* The program is therefore first to define a general random field (R.F.), then a Gibbs ensemble or Gibbs random field (G.R.F.), and next a Markov random field (M.R.F.). The easy half of the theorem will be the statement that every G.R.F. is a M.R.F. This is due to the simple explicit form of the definition of a G.R.F. The converse, that every M.R.F. is a G.R.F., is less

---

Frank Spitzer received his Michigan PhD under D. Darling. He held positions at Cal Tech and the University of Minnesota before his present post at Cornell. He spent leaves of absence at Princeton and at the Univ. de Strasbourg (Senior NSF and Guggenheim Fellowships). His extensive published work in probability theory includes *Principles of Random Walk* (Van Nostrand, 1964). *Editor.*

obvious, and perhaps surprising since the usual derivation of the Gibbs formula defining a G.R.F. depends on the physical fact that the interaction between particles is described by a potential. No such notions from physics enter into the purely probabilistic definition of a M.R.F.

Added in proof: Almost identical results have been obtained by M. B. Averintzev, "On a method of describing discrete parameter random fields," *Problemy Peredači Informacii*, Vol. 6, no. 2, 1970, pp. 100–109.

2. Basic Definitions.

Definition of a random field. Let Z^ν denote the ν-dimensional integers, or the lattice points in ν-dimensional space. For x, y in Z^ν, $|x-y|$ is the Euclidean distance from x to y. A finite subset $D\subset Z^\nu$ is called a domain if it is connected, i.e., if x, $y\in D$ implies the existence of a path $x=x_0, x_1, \cdots, x_{n-1}, x_n=y$, such that all $x_i\in D$ and $|x_i-x_{i+1}|=1$. The boundary ∂D of a domain D consists of all y in $Z^\nu-D$ which have one or more neighbors in D. (x and y are neighbors if $|x-y|=1$.) Finally we denote $\bar{D}=D\cup\partial D$. If $\Omega=\{0, 1\}^D$, if $\mathfrak{F}$ is the collection of all subsets of Ω and if P is a probability measure on $\mathfrak{F}$, then the triple $(\Omega, \mathfrak{F}, P)$ *is* called *a random field* (R.F.) *on the domain D*. Thus we may think of a R.F. on D as a probability measure on the set of all maps $\omega: D\to\{0, 1\}$, and consequently as a probability measure on the set of all possible configurations of particles on D. (We think of a site $x\in D$ as occupied by a particle if $\omega(x)=1$, and as empty if $\omega(x)=0$.) Thus the configuration of particles described by $\omega\in\Omega$ is the set $\{x: x\in D, \omega(x)=1\}$. In the one dimensional case ($\nu=1$) we can also think of the R.F. on a domain (interval) D as a stochastic process $\omega(x)$ with time parameter $x\in D$, and values in the two point set $\{0, 1\}$.

Definition of a Gibbs random field (G.R.F.). A function U from $Z^\nu\times Z^\nu$ to $\mathfrak{R}$ (the reals) is called a symmetric homogeneous, nearest neighbor pair potential (or briefly *pair potential* from now on) if for all x, y in Z^ν

(i) $U(x, y)=U(y, x)$ (symmetry),

(ii) $U(x, y)=U(0, y-x)$ (homogeneity),

(iii) $U(x, y)=0$ when $|y-x|>1$ (nearest neighbor property).

Before we can define the most general G.R.F. with pair potential U we also have to specify a boundary value (B.V.) function. This is an arbitrary map $\phi:\partial D\to\{0, 1\}$. When ϕ is given it will be convenient to extend each $\omega\in\Omega$ to a map $\bar{\omega}$ of $\bar{D}\to\{0, 1\}$ by the rule

$$\bar{\omega}(x) = \begin{cases}\omega(x) & \text{for } x\in D,\\ \phi(x) & \text{for } x\in\partial D.\end{cases}$$

Suppose we are given a domain $D\subset Z^\nu$, and a potential U satisfying (i), (ii), (iii), and a B.V. function ϕ. Then we shall say that *a* R.F. $(\Omega, \mathfrak{F}, P)$ *on D is a* G.R.F. *with pair potential U and* B.V. *function* ϕ, *if P is defined by the Gibbs formula*

$$P(\omega) = Z^{-1}\exp\left[-\frac{1}{2}\sum_{x\in\bar{D}}\sum_{y\in\bar{D}}\bar{\omega}(x)\bar{\omega}(y)U(x, y)\right], \qquad \omega\in\Omega. \tag{1}$$

Here Z is the unique normalizing constant for which

(2) $$\sum_{\omega\in\Omega} P(\omega) = 1.$$

In particular, if the B.V. function $\phi\equiv 0$ on ∂D, then we get the G.R.F. with B.V. zero, given by

(3) $$P(\omega) = Z^{-1}\exp\left[-\frac{1}{2}\sum_{x\in D}\sum_{y\in D}\omega(x)\omega(y)U(x,y)\right], \qquad \omega\in\Omega.$$

There is another interesting possibility. If D happens to be a rectangle then we can identify opposite points to produce a lattice *torus T without boundary.* In this case the G.R.F. on T is called a *periodic* G.R.F., and its probability measure P is defined by formula (3) with D replaced by T.

Definition of a Markov random field (M.R.F.). This definition is more intuitive but less explicit than that of a G.R.F. As in the case of the G.R.F. we assume given a B.V. function $\phi:\partial D\to\{0, 1\}$ and shall define different M.R.F.'s on a domain D corresponding to different B.V. functions ϕ. When D is replaced by a torus T, the B.V. function becomes unnecessary and we shall define a periodic M.R.F. A R.F. $(\Omega, \mathfrak{F}, P)$ on D will be called a M.R.F. if it satisfies the three conditions (a), (b), (c) below. First

(a) $$P(\omega) > 0 \qquad \text{for each } \omega\in\Omega \text{ (positivity).}$$

In view of (a) we can define the *one point conditional probabilities*

(4) $$P[\omega(x) = 1 \mid \bar{\omega}(\cdot) = f(\cdot) \text{ on } \overline{D} - \{x\}], \qquad x\in D,$$

by the elementary formula $P(A\mid B) = P(AB)/P(B)$. Let us clarify (4) which is written in dangerously brief notation. The map $f:\overline{D}-\{x\}\to\{0, 1\}$ is quite arbitrary except that it must agree with the B.V. function ϕ on ∂D (unless of course $D=T$ in which case there is no boundary). Thus (4) represents the probability that $\omega(x)=1$ (that there is a particle at x), given that $\omega(y)=f(y)$ at the points y of $D-\{x\}$, and given, in addition, the boundary values $f(z)=\phi(z)$ for $z\in\partial D$. The latter are deterministic and will therefore be treated as events of probability one. Now the second condition defining a M.R.F. may be stated as

(b) the conditional probabilities in (4) depend only on the values of f at the points y in $\overline{D}$ such that $|y-x|=1$ (nearest neighbor condition).

The third and last defining condition for a M.R.F. is

(c) the conditional probabilities in (4) are translation invariant, i.e., x, y in D implies

$$P[\omega(x) = 1 \mid \bar{\omega}(\cdot) = f(\cdot) \text{ on } \overline{D} - \{x\}] = P[\omega(y) = 1 \mid \bar{\omega}(\cdot) = g(\cdot) \text{ on } \overline{D} - \{y\}],$$

whenever $f(x+z)=g(y+z)$ for all z with $|z|=1$ (homogeneity).

Note that we have not yet proved the existence of a R.F. which satisfies (a), (b), and (c). Nevertheless, *if* $(\Omega, \mathfrak{F}, P)$ *is a* R.F. *on a domain D which satisfies*

(a), (b), (c), *then we shall say that* $(\Omega, \mathfrak{F}, P)$ *is a* M.R.F. *on* D *with* B.V. *function* ϕ (*or a periodic* M.R.F. *on* $D=T$ *when* D *is made into a torus* T *without boundary*).

3. The Main Theorem.

MAIN THEOREM: *Every M.R.F. on a domain D with B.V. function ϕ is a G.R.F. on D with B.V. function ϕ and vice versa. The same statement holds for periodic random fields. The explicit correspondence between the conditional probabilities of the M.R.F. and the pair potential of the corresponding G.R.F. is given by equations* (5), (6), (8), *and* (9) *below.*

In the proof we shall work on a fixed domain $D \subset Z^\nu$ with a fixed B.V. function ϕ and ignore the periodic case which can be handled by the same method. Step 1 of the proof will show that every G.R.F. is a M.R.F. Step 2 will then show that for every M.R.F. there exists a G.R.F. with the same conditional probabilities as the M.R.F. Finally the last step, step 3, will show that there exists at most one R.F. satisfying (a), (b), (c), with given conditional probabilities. It will be apparent that this completes the proof of the main theorem.

STEP 1: (*Every* G.R.F. *is a* M.R.F.). We start with a G.R.F. whose probability measure P is given by (1) and proceed to verify (a), (b), and (c). Condition (a) is obvious since exponentials are positive. To check (b) and (c) we compute the one point conditional probabilities

$$P[\omega(x) = 1 \mid \bar{\omega} = f \text{ on } \bar{D} - \{x\}] = \frac{P[\omega(x) = 1 \text{ and } \bar{\omega} = f]}{P[\omega(x) = 1 \text{ and } \bar{\omega} = f] + P[\omega(x) = 0 \text{ and } \bar{\omega} = f]}.$$

According to (1) the probability in the numerator is

$$Z^{-1} \exp\left\{ -\frac{1}{2}\left[\sum_{s \in \bar{D}-\{x\}} \sum_{t \in \bar{D}-\{x\}} f(s)f(t)\, U(s,t) + U(x,x) + 2 \sum_{s \in \bar{D}-\{x\}} f(s) U(s,x) \right]\right\},$$

while the second probability in the denominator is

$$Z^{-1} \exp\left[-\frac{1}{2} \sum_{s \in \bar{D}-\{x\}} \sum_{t \in \bar{D}-\{x\}} f(s) f(t)\, U(s,t) \right].$$

A brief calculation therefore gives, for each map $f: \bar{D} \to \{0, 1\}$ such that $f = \phi$ on ∂D,

$$P[\omega(x) = 1 \mid \bar{\omega} = f \text{ on } \bar{D} - \{x\}] = \frac{1}{1 + \exp\left[\dfrac{1}{2} U(x,x) + \sum_{s \in \bar{D}-\{x\}} f(s) U(s,x)\right]}. \tag{5}$$

Let

$$U(0, 0) = u_0,\ U(0, l_k) = u_k, \qquad 1 \leqq k \leqq \nu,$$

where l_k are the unit vectors with kth component $+1$ and all other components zero. Then the pair potential U is uniquely determined by these $\nu+1$ parameters, and properties (ii) and (iii) further imply that (5) takes the form

$$P[\omega(x) = 1 \mid \bar{\omega} = f \text{ on } \bar{D} - \{x\}]$$

$$= \frac{1}{1 + \exp\left\{(u_0/2) + \sum_{k=1}^{\nu} [f(x + l_k) + f(x - l_k)]u_k\right\}}. \tag{6}$$

Clearly the right hand side of (6) exhibits properties (b) and (c), i.e., the nearest neighbor property and translation invariance of the conditional probabilities. Therefore every G.R.F. is a M.R.F. with the same B.V. function (or without one in the periodic case).

STEP 2: (*Existence of a G.R.F. with the same conditional probabilities as a given M.R.F.*). While a G.R.F. in Z^ν is determined by $\nu+1$ real parameters (the constants $u_0, u_1, \cdots, u_\nu$ in the last section) it is not at all clear "how many" different M.R.F.'s there are. The conditional probabilities must clearly satisfy certain consistency conditions which reduce the number of possibilities. Indeed the key result of this section is that in Z^ν there is a $\nu+1$ parameter family of possible conditional probability functions. To make this precise let $(\Omega, \mathfrak{F}, P)$ be a given M.R.F. on $D \subset Z^\nu$ and introduce the conditional probabilities

$$\begin{aligned} p_0 &= P[\omega(x) = 1 \mid \bar{\omega}(x + y) = 1, \text{ whenever } |y| = 1], \\ p_k &= P[\omega(x) = 1 \mid \bar{\omega}(x + l_k) = 0 \text{ and } \bar{\omega}(x + y) = 1 \text{ for all} \\ &\qquad \text{other } y \text{ with } |y| = 1], \qquad 1 \leq k \leq \nu. \end{aligned} \tag{7}$$

Let further p_k' be defined just as p_k except that l_k is replaced by $-l_k$ in the definition of p_k. Then we have the following:

CONSISTENCY LEMMA. *Let $(\Omega, \mathfrak{F}, P)$ be a M.R.F. on a sufficiently large domain $D \subset Z^\nu$. Then all the conditional probabilities in* (4) *are uniquely determined by the $\nu+1$ parameters $p_0, p_1, \cdots, p_\nu$. In particular we have $p_k' = p_k$, for $1 \leq k \leq \nu$.*

According to this lemma it will suffice to construct a G.R.F. with the same conditional probabilities $(p_0, p_1, \cdots, p_\nu)$ as the given M.R.F. If we had such a G.R.F. then we could assert, in view of (6), that

$$\begin{aligned} p_0 &= \frac{1}{1 + \exp\left[(u_0/2) + 2\sum_{j=1}^{\nu} u_j\right]}, \\ p_k &= \frac{1}{1 + \exp\left[(u_0/2) + 2\sum_{j=1}^{\nu} u_j - u_k\right]}, \qquad 1 \leq k \leq \nu. \end{aligned} \tag{8}$$

Here $(p_0, p_1, \cdots, p_\nu)$ is the set of parameters of the given M.R.F. as defined

by (7), and $(u_0, u_1, \cdots, u_\nu)$ specifies the potential of the G.R.F., since $u_0 = U(x, x)$, $u_k = U(x, x+l_k) = U(x, x-l_k)$ for $1 \leqq k \leqq \nu$. Observe now that (8) maps the $\nu+1$ dimensional Euclidean space $-\infty < u_k < \infty$, $0 \leqq k \leqq \nu$, in a 1:1 manner onto the $\nu+1$ dimensional cube $0 < p_k < 1$, $0 \leqq k \leqq \nu$. If we introduce the auxiliary parameters

$$\alpha_k = \log(p_k^{-1} - 1), \qquad 0 \leqq k \leqq n,$$

then one can in fact invert (8) explicitly to obtain

$$(9) \qquad \begin{aligned} u_0 &= 4 \sum_{j=1}^{\nu} \alpha_j - (4\nu - 2)\alpha_0, \\ u_k &= \alpha_0 - \alpha_k, \qquad 1 \leqq k \leqq \nu. \end{aligned}$$

It follows then that, given a M.R.F. which is determined by $(p_0, p_1, \cdots, p_\nu)$, we obtain a G.R.F. with the same conditional probabilities by choosing the potential $U(x, y)$ which is determined by the values $(u_0, u_1, \cdots, u_\nu)$ in (9).

Proof of the consistency lemma: We begin by introducing, *ad hoc*, certain elementary identities valid for a *completely arbitrary* probability space $(\Omega, \mathfrak{F}, P)$. Let A, B, C denote three arbitrary events such that all possible intersections of A, B, C and of their complements $\bar{A}$, $\bar{B}$, $\bar{C}$ have strictly positive probabilities (excluding of course the empty sets $A\bar{A}$, etc.). (We write $\bar{A}$ for the complement of A, apologizing for the previous use of $\bar{D}$ to denote $D \cup \partial D$, and AB for $A \cap B$.) It follows that all conditional probabilities of the form $P(A \mid BC)$, $P(AB \mid C)$, $P(A \mid \bar{B}C)$, etc., are well defined and strictly positive. We assert that

$$(10) \qquad \begin{aligned} \frac{1}{P(AB \mid C)} &= \frac{1}{P(A \mid BC)} + \frac{1}{P(B \mid AC)P(A \mid \bar{B}C)} - \frac{1}{P(A \mid \bar{B}C)} \\ &= \frac{1}{P(B \mid AC)} + \frac{1}{P(A \mid BC)P(B \mid \bar{A}C)} - \frac{1}{P(B \mid \bar{A}C)}. \end{aligned}$$

The proof of the first part of (10) is immediate by substitution of the definitions $P(AB \mid C) = P(ABC)/P(C)$, etc. The second half follows from the first by observing that $P(AB \mid C)$ depends symmetrically on A and B. It will further be convenient to work with the functions $H(\cdot \mid \cdot)$ defined by

$$H(A \mid B) = \frac{1}{P(A \mid B)} - 1.$$

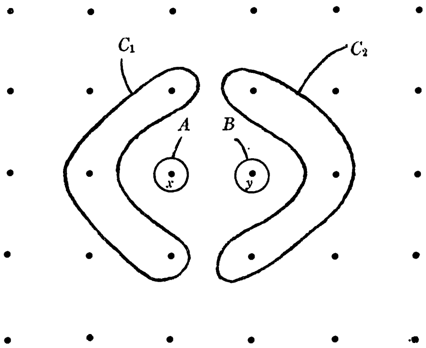
Direct substitution then reduces (10) to the simple form

$$(11) \qquad H(B \mid AC)H(A \mid \bar{B}C) = H(A \mid BC)H(B \mid AC).$$

Returning to the proof of the consistency lemma, we assume that $(\Omega, \mathfrak{F}, P)$ is a given M.R.F. on a domain $D \subset Z^\nu$, and we fix two neighboring points x and y in D. Then we define the following events A, B, C_1, C_2, and C in $\mathfrak{F}$. Let A be the event that $\omega(x) = 1$, so that $\bar{A} = \{\omega : \omega(x) = 0\}$. Similarly $B = \{\omega : \omega(y) = 1\}$.

Next C_1 is a description of the values of $\bar{\omega}(\cdot)$ on all the 2ν neighbors of x *except* at the point y. For instance, if $\nu=2$ and $y=x-l_2$, then C_1 might be chosen as the event $\{\omega:\bar{\omega}(x+l_1)=1,\ \bar{\omega}(x-l_1)=0,\ \bar{\omega}(x+l_2)=1\}$. Note that we have expressed C_1 in terms of $\bar{\omega}$ instead of ω since some of the neighbors of x may be points of the boundary ∂D. Similarly C_2 will be a description of $\bar{\omega}(\cdot)$ on the 2ν neighbors of y except x. Finally we define C as the intersection $C=C_1C_2$.

Consider now the conditional probability $P(B|AC_2)$, i.e., the probability that $\omega(y)=1$, given that $\omega(x)=1$, and that the values of $\bar{\omega}$ at the remaining neighbors of y are described by C_2. In view of property (b) of the conditional probabilities of a M.R.F. there is no new information about $\omega(y)$ conveyed by specifying C_1 in addition to A and C_2 (cf. the drawing for dimension $\nu=2$).



$$\Rightarrow \quad \begin{aligned} P(B|AC_2) &= P(B|AC_1C_2) = P(B|AC), \\ P(A|BC_1) &= P(A|BC_1C_2) = P(A|BC). \end{aligned}$$

Therefore $P(B|AC_2)=P(B|AC)$, $P(A|BC_1)=P(A|BC)$. Also replacing A by $\bar{A}$ in the first identity and B by $\bar{B}$ in the second, $P(B|\bar{A}C_2)=P(B|\bar{A}C)$, $P(A|\bar{B}C_1)=P(A|\bar{B}C)$. From the definition of $H(\cdot|\cdot)$ it then follows that

$$\begin{aligned} H(B|AC_2) &= H(B|AC), & H(A|BC_1) &= H(A|BC), \\ H(B|AC_2) &= H(B|AC), & H(A|\bar{B}C_1) &= H(A|\bar{B}C). \end{aligned}$$

Now substitution into (11) yields

$$H(B|AC_2)H(A|\bar{B}C_1) = H(A|BC_1)H(B|AC_2). \tag{12}$$

The consistency lemma readily follows from (12) by making suitable choices for the events C_1 and C_2. First let C_1 be the event that $\bar{\omega}(z)=1$ at all the neighbors z of x except y, and C_2 the event that $\bar{\omega}(z)=1$ at all the neighbors z of y except x. Recalling the definition of p_0, p_k, p_k' with $1\leqq k\leqq\nu$, it is clear that (12) reads (in the case when $y=x+l_k$)

$$\left(\frac{1}{p_0}-1\right)\left(\frac{1}{p_k}-1\right) = \left(\frac{1}{p_0}-1\right)\left(\frac{1}{p_k'}-1\right). \tag{13}$$

By proper choice of x and y we show that (13) holds for all $1\leqq k\leqq\nu$. Hence we have

$$p_k' = p_k, \qquad \text{for } 1 \leqq k \leqq \nu. \tag{14}$$

REMARK: There is the unpleasant possibility that the domain D is too small to contain pairs of neighbors x and $y=x+l_k$ for each k, $1 \leqq k \leqq \nu$. In this case the conclusion of the consistency lemma is false. (That is why D was required to be large enough in its statement.) If D is too small, then only a subset of the conditional probabilities are determined by $(p_0, p_1, \cdots, p_\nu)$. But it can be checked that these are the only parameters needed to describe the corresponding G.R.F. Thus the conclusion of step 2 remains correct even when the domain D is too small for the consistency lemma to hold.

The rest of the proof of the consistency lemma proceeds by induction. Suppose we have shown that the parameters $(p_0, p_1, \cdots, p_\nu)$ determine uniquely all conditional probabilities of the form $P[\omega(x)=1 \mid \bar{\omega}=f \text{ on } \overline{D}-\{x\}]$ *where* $f=1$ *on all but at most* j *of the* 2ν *neighbors of* x. Fix the point x, and let C be a description of $\bar{\omega}=f$ on the neighbors of x such that $f(y)=0$ on exactly $j+1$ of the neighbors y of x. We then have to show that $P[\omega(x)=1 \mid C]$ is uniquely determined by $(p_0, p_1, \cdots, p_\nu)$. Suppose now that $f(y)=0$ for $y=x+l_k$. (If this is not the case for any k then $f(y)=0$ for some y of the form $x-l_k$ and the reasoning which follows will apply without change.) Let C' be the modification of C obtained by changing the value of f at $y=x+l_k$ from 0 to 1. Let $A=\{\omega:\omega(x)=1\}$, $B=\{\omega:\omega(y)=1\}$. Let C_1 be the event C with the specification at $y=x+l_k$ omitted, so that $C=C_1\overline{B}$ and $C'=C_1B$. Finally let C_2 be the event that $\bar{\omega}(\cdot)=1$ at all the neighbors of y except x. Now consider equation (12) which after substitution of $C=C_1\overline{B}$, $C'=C_1B$ becomes

$$H(B \mid AC_2)H(A \mid C) = H(B \mid AC_2)H(A \mid C'). \tag{15}$$

We shall show that $H(A \mid C)$, and hence $P(A \mid C)$ is uniquely determined by the parameters $(p_0, p_1, \cdots, p_\nu)$. This will follow if the other three terms in (15) are so determined (note that $H(B \mid AC_2)>0$). Now the configuration described by C' contains exactly j zeros, that described by $\overline{A}C_2$ contains exactly one zero, and that described by AC_2 no zero at all. Thus it follows from the induction hypothesis that $H(A \mid C')$ is determined by $(p_0, \cdots, p_\nu)$, and from the fact that $p_k'=p_k$ (already proved) that $H(B \mid \overline{A}C_2)$ is determined (since $H(B \mid \overline{A}C_2)$ is either $p_k^{-1}-1$ or $(p_k')^{-1}-1$). That completes the induction step from j to $j+1$, and hence the proof of the consistency lemma, which was already shown to complete step two of the proof of the main theorem.

STEP 3 (*Every* M.R.F. *is a* G.R.F.). Let us review the logic of steps one and two. According to step two there exists, for every given M.R.F., a G.R.F. with the same conditional probabilities. According to step one this G.R.F. is also a M.R.F. But we have *not yet shown that this is the same* M.R.F. *as the given* M.R.F. *we started with*. Thus we have to show that there exists only one R.F. with the same one point conditional probabilities as a given M.R.F. Actually we shall do much more. Let $(\Omega, \mathfrak{F}, P)$ be an arbitrary *positive* R.F., i.e., a R.F. such that $P(\omega)>0$ for each $\omega \in \Omega$. We shall show that *every positive* R.F. *is uniquely deter-*

mined by its conditional probabilities of the form $P[\omega(x)=1 \mid \bar{\omega}(\cdot)=f(\cdot)$ *on* $\bar{D}-\{x\}]$ *for all possible choices of* $x \in D$ *and* $f: \bar{D}-\{x\} \to \{0, 1\}$. But this assertion in turn can be generalized and, in the process, simplified. Let $n=|D|$, the cardinality of D, let $\{x_1, x_2, \cdots, x_n\}$ be an enumeration of D, and define $A_n=\{\omega : \omega(x_n)=1\}$. Then $\mathfrak{F}$ is the algebra of subsets of Ω generated by $A_1, A_2, \cdots, A_n$. Let $\mathcal{G}_k$ be the subset of $\mathfrak{F}$ consisting of all events of the form

$$A = \bigcap_{\substack{1 \leq i \leq n \\ i \neq k}} B_i, \qquad \text{where each } B_i = A_i \quad \text{or} \quad A_i. \tag{16}$$

Then the one point conditional probabilities of the R.F. $(\Omega, \mathfrak{F}, P)$ are all the probabilities of the form

$$P(A_k \mid A), \qquad A \in \mathcal{G}_k, \quad 1 \leqq k \leqq n. \tag{17}$$

The boundary values are of course thought of as included in A when needed, i.e., when x_k has neighbors in ∂D; they cause no trouble since they are given with probability one. Our assertion now becomes that the probability measure P on $(\Omega, \mathfrak{F})$ is uniquely determined by the probabilities in (17). This fact can be reformulated in general terms, without any reference to random fields.

LEMMA ON CONDITIONAL PROBABILITIES. *Let Ω be an arbitrary set with n given subsets $A_1, A_2, \cdots, A_n$. Let $\mathfrak{F}$ be the algebra of subsets of Ω generated by $A_1, \cdots, A_n$. Let P be a probability measure on $(\Omega, \mathfrak{F})$ such that $P(C)>0$ for all C in $\mathfrak{F}$ except the empty set (so that all possible conditional probabilities $P(A \mid B)=P(AB)/P(B)$ are defined when $B \neq \varnothing$). Let $\mathcal{G}_k \subset \mathfrak{F}$ be the set of events of the form* (16), *and suppose we know all the conditional probabilities of* (17). *Then these conditional probabilities determine the probability measure P uniquely.*

Proof: Let S denote a subset of $N=\{1, 2, \cdots, n\}$ and let

$$p(S) = P\Big[\bigcap_{i \in S} A_i \cap \bigcap_{j \notin S} A_j\Big].$$

Then P will be completely determined on $\mathfrak{F}$ if $p(S)$ is known for every $S \subset N$. But the conditional probabilities in (17) can be written

$$P(A_k \mid A) = \frac{P(A_k \cap A)}{P(A)} = \frac{p(S \cup \{k\})}{p(S \cup \{k\}) + p(S)} = \left[1 + \frac{p(S)}{p(S \cup \{k\})}\right]^{-1}, \tag{18}$$

if we choose S to correspond to A in such a way that

$$A = \bigcap_{i \in S} A_i \cap \bigcap_{j \in N-(S \cup \{k\})} A_j.$$

It follows from (18) that the conditional probabilities in (17) determine $p(S \cup \{k\})/p(S)$ for every $S \subset N$, and $k \in N$ such that $k \notin S$. Now let $S=\{i_1, i_2, \cdots, i_\nu\}$ and observe that

$$(19)\qquad \frac{p(S)}{p(\emptyset)} = \frac{p(\{i_1\})}{p(\emptyset)} \frac{p(\{i_1, i_2\})}{p(\{i_1\})} \cdots \frac{p(\{i_1, \cdots, i_\nu\})}{p(\{i_1, \cdots, i_{\nu-1}\})}$$

is then determined for every $S \subset N$. Summing (19) over all $S \subset N$ determines $p(\emptyset)$, since $\sum p(S) = 1$ and reapplying (19) shows that all $p(S)$ are determined. That completes the proof of the lemma, and hence of the main theorem.

4. Examples of special Markov random fields.

I. ROTATION INVARIANT FIELDS. If we require that a G.R.F. be rotation invariant, then this means that the potential U defining it is rotation invariant. Hence the number of parameters is reduced to two, since we obtain

$$U(x, x) = u_0, \qquad U(x, x + l_k) = u_k = u_1 \qquad \text{for } 1 \leqq k \leqq \nu.$$

Looking at the same R.F. as a M.R.F. then it follows from (8) that the conditional probabilities $(p_0, p_1, \cdots, p_\nu)$ satisfy

$$(20)\qquad \begin{aligned} p_0^{-1} - 1 &= \exp\left[\frac{u_0}{2} + 2\nu u_1\right], \\ p_k^{-1} - 1 &= \exp\left[\frac{u_0}{2} + (2\nu - 1)u_1\right], \qquad 1 \leqq k \leqq \nu. \end{aligned}$$

It is easy to see, using (5), that all other conditional probabilities must then be rotation invariant as well. Thus the family of all rotation invariant M.R.F.'s on a domain $D \subset Z^\nu$ is described by *two parameters* (p_0, p_1) *regardless of the dimension* ν. This fact was far from obvious from the definition of a M.R.F.

II. FIELDS WITHOUT INTERACTION. Let us define lack of interaction in a G.R.F. by saying that the pair potential $U(x, y)$ vanishes unless $x = y$. This means that u_0 is arbitrary while $u_1 = u_2 = \cdots = u_\nu = 0$. In view of (8) the conditional probabilities then satisfy $p_0 = p_1 = p_2 = \cdots = p_\nu$, and going back to the Gibbs formula (5) shows that

$$P[\omega(x) = 1 \mid \bar{\omega} = f \text{ on } \bar{D} - \{x\}] = p_0,$$

independently of f. Thus lack of interaction in the sense of a trivial pair potential is equivalent to each conditional probability being independent of its condition. Equivalently, this means that the family of *random variables* $\omega(x)$, $x \in D$, *are mutually independent.*

III. ONE-DIMENSIONAL FIELDS. Let

$$(21)\qquad M = \begin{pmatrix} M_{00} & M_{01} \\ M_{10} & M_{11} \end{pmatrix} = \begin{pmatrix} p & 1-p \\ 1-q & q \end{pmatrix}, \qquad 0 < p < 1, \quad 0 < q < 1,$$

be the transition matrix of a Markov chain $\{\xi(n)\}$, $n \geqq 0$, whose state space is the two point set $\{0, 1\}$. Fix a positive integer r, and let $D = \{1, 2, \cdots, r\} \subset Z^1$. Let ϕ be a given B.V. function $\phi: \partial D \to \{0, 1\}$. Let $\{\eta(1), \eta(2), \cdots, \eta(r)\}$ de-

note a family of random variables, with values in $\{0, 1\}$, whose joint distribution is the same as the joint conditional distribution of the set $\{\xi(1), \cdots, \xi(r)\}$ when it is subject to the condition $\xi(0)=\phi(0)$, $\xi(r+1)=\phi(r+1)$. Then it is not hard to show that *the collection* $\{\eta(x), x\in D\}$ *is a* M.R.F. *on* D *with* B.V. *function* ϕ. Indeed one obtains the class of all possible M.R.F.'s with B.V. function ϕ on D in the same way, by varying the parameters p, q in the definition of the matrix M in (21). It can be shown that the explicit correspondence between (p, q) in (21) and the conditional probabilities (p_0, p_1) of the M.R.F. is given by

$$\frac{1}{p_0} - 1 = \frac{(1-p)(1-q)}{q^2}, \quad \frac{1}{p_1} - 1 = \frac{p}{q}, \tag{22}$$

which is a 1:1 map of the square $\{0<p<1, 0<q<1\}$ onto the square $\{0<p_0<1, 0<p_1<1\}$.

IV. Symmetric random fields. We want a M.R.F. or G.R.F. which is rotation invariant and possesses the additional symmetry property of *invariance under interchange of the two symbols* 0 *and* 1. In other words, let $\pi_k(1)$ denote the probability that $\omega(x)=1$ given that $\bar{\omega}(\cdot)=1$ at exactly k of the 2ν neighbors of x. Define $\pi_k(0)$ in the same way, but with 0 and 1 interchanged. Then we require that $\pi_k(0)=\pi_k(1)$ for all $0\leqq k\leqq 2\nu$. This will be the case if and only if

$$u_0 + 2\nu u_1 = 0, \quad \text{or}$$

$$\frac{1}{p_0} - 1 = \exp\left[-\frac{u_0}{2}\right], \quad \frac{1}{p_1} - 1 = \exp\left[-\frac{u_0}{2}\left(1-\frac{1}{\nu}\right)\right]. \tag{23}$$

Proof: It follows from (6) that

$$[\pi_k(1)]^{-1} - 1 = \exp\left[\frac{u_0}{2} + ku_1\right],$$

$$[1-\pi_k(0)]^{-1} - 1 = \exp\left[\frac{u_0}{2} + (2\nu - k)u_1\right].$$

This implies that $\pi_k(0)=\pi_k(1)$ for all $0\leqq k\leqq 2\nu$ if and only if $u_0+2\nu u_1=0$.

5 Infinite random fields. Some brief remarks will place this note in the context of interesting recent work on the mathematics and physics of infinite random fields. The latter are of physical interest because the thermodynamic description of a lattice gas in Z^ν is only obtained after passage to the limit from a finite rectangle $D\subset Z^\nu$ to all of Z^ν. (See [3] for a systematic treatment of the theory of this so-called thermodynamic limit.)

To define an infinite R.F., or a R.F. on all of Z^ν, let $\Omega=\{0, 1\}^{Z^\nu}$, and take for $\mathfrak{F}$ the smallest Borel field of subsets of Ω which contains the class $\mathcal{C}$ of all cylinder sets of the form

$$A = [\omega \mid \omega(x) = \epsilon(x) \text{ for } x \in D],$$

for all finite $D \subset Z^\nu$ and all maps $\epsilon: D \to \{0, 1\}$. Finally let P be a probability measure on $(\Omega, \mathfrak{F})$. Then the probability triple $(\Omega, \mathfrak{F}, P)$ is called a R.F. on Z^ν.

Next, let us say that a R.F. $(\Omega, \mathfrak{F}, P)$ on Z^ν is an infinite M.R.F. if it satisfies (a′), (b′), (c′) below. First,

(a′) $$P(C) > 0 \qquad \text{for all } C \in \mathfrak{C}.$$

If (a′) holds then the conditional probabilities $P[\omega(x) = 1 \mid \omega = f \text{ on } D - \{x\}]$ are well defined for every finite set $D \subset Z^\nu$ such that D contains x and also all its 2ν neighbors. Thus it makes sense to require

(b′) the above conditional probabilities depend only on the values of f at the neighbors of x,

(c′) the above conditional probabilities are translation invariant.

It follows from step 2 of our proof of the main theorem for finite random fields that *even for an infinite* M.R.F. *the conditional probabilities may be assumed to be given by* (5) *or* (6). Note however that the definition of an infinite R.F. by the Gibbs formula (3) is impossible, since $P(\omega) = 0$ for each ω (the set Ω being uncountable). This led Dobrushin [2] to define an infinite G.R.F. on Z^ν by the requirement that it be a M.R.F. with conditional probabilities given by (6), and to study the following basic questions:

(I) Does there exist such a R.F. for every possible set of parameter values $(u_0, u_1, \cdots, u_\nu)$?

(II) If so, is it unique?

The answer to the existence question (I) is YES. (See Dobrushin [1], Theorem 1.) The exciting answer to the uniqueness question (II), on the other hand, is SOMETIMES, i.e., there is uniqueness for certain but not for all parameter sets $(u_0, u_1, \cdots, u_\nu)$.

For a more detailed answer to II, consider first the one dimensional case. Then there is a unique infinite G.R.F. with given (u_0, u_1), and it is not hard to construct it from the Markov chain on $\{0, 1\}$ with transition matrix M in (21). There is a unique strictly stationary process $\{\zeta(n)\}$, $-\infty < n < \infty$, with state space $\{0, 1\}$, such that $\{\zeta(n)\}$ for $n \geqq 0$, conditioned on $\zeta(0) = 0$ (or 1) is a Markov chain with transition matrix M and initial state 0 (or 1). It is not hard to show that this process $\{\zeta(n)\}$, $-\infty < n < \infty$, is the unique G.R.F. on Z^1 with parameters (u_0, u_1) if (p, q) in (21) is chosen according to (22). Indeed this uniqueness is a special case of [1], Theorem 3.

Exercise: Show that the unique one dimensional infinite M.R.F. with $U(x, x) = u_0$, $U(x, x \pm l_k) = u_1$ has the particle density

$$\rho = P[\omega(x) = 1] = \left[\frac{1-q}{1-p} + 1\right]^{-1} = \frac{1}{2}\left[1 + \frac{h_1 - 1}{\sqrt{(h_1 - 1)^2 + 4h_0}}\right], \qquad x \in Z^1,$$

where (p, q) are the parameters in (21), and

$$h_0 = p_0^{-1} - 1 = \exp\left[\frac{u_0}{2} + 2u_1\right], \qquad h_1 = p_1^{-1} - 1 = \exp\left[\frac{u_0}{2} + u_1\right].$$

In dimension $\nu \geqq 2$ the situation is quite different. Uniqueness can be shown only for certain values of the parameters $(u_0, u_1, \cdots, u_\nu)$, for example by use of Theorem 2 of [1]. The simplest known examples of nonuniqueness are obtained for the symmetric random fields in example IV of the last section. In other words, let $\nu \geqq 2$ and assume that $u_0 + 2\nu u_k = 0$ for $1 \leqq k \leqq \nu$, so that (23) holds. Then the uniqueness question becomes:

Does there exist a unique infinite M.R.F. (*or* G.R.F.) *on* Z^ν, $\nu \geqq 2$ *with conditional probabilities*

$$p_0 = \frac{1}{1 + \exp[-(u_0/2)]},$$
$$p_1 = p_2 = \cdots = p_\nu = \frac{1}{1 + \exp[-(u_0/2)(1 - (1/\nu))]}? \tag{24}$$

The answer is *yes, if* u_0 *is sufficiently small.* On the other hand it is *no, for all sufficiently large* u_0 (see [2], pp. 306–308). In the physical description of a gas by such a R.F. the parameter u_0 is inversely proportional to the *absolute temperature* T. The existence of two infinite M.R.F.'s satisfying (24) for all large u_0 can be interpreted as the coexistence of two distinct phases (gas and liquid) of a substance when the temperature T is sufficiently low.

For an indication of the proof of nonuniqueness, let D_N be a sequence of cubes of side length N in Z^ν, $\nu \geqq 2$. Let $\{\omega_N^{(0)}(x), x \in D_N\}$ and $\{\omega_N^{(1)}(x), x \in D_N\}$ be finite symmetric random fields on $D_N \subset Z^\nu$, the former with B.V. function $\phi \equiv 0$ on ∂D_N, the latter with B.V. $\phi \equiv 1$ on ∂D_N. By combinatorial arguments, made possible by the symmetry of these fields, it is shown in [2] that for sufficiently large u_0 there exists a constant $\gamma < \frac{1}{2}$ such that

$$P[\omega_N^{(0)}(x) = 1] \leqq \gamma < \tfrac{1}{2}, \qquad P[\omega_N^{(1)}(x) = 1] \geqq 1 - \gamma > \tfrac{1}{2},$$

for all $x \in D_N$, and every positive integer N. This suggests the existence of two infinite G.R.F.'s, a low density "gas" $\{\omega^{(0)}(x), x \in Z^\nu\}$, and a high density "liquid" $\{\omega^{(1)}(x), x \in Z^\nu\}$, which have particle densities

$$\rho_0 = P[\omega^{(0)}(x) = 1] \leqq \gamma, \qquad \rho_1 = P[\omega^{(1)}(x) = 1] \geqq 1 - \gamma, \quad x \in Z^\nu,$$

and which have both the same conditional probabilities $(p_0, p_1, \cdots, p_\nu)$, given by (24) in terms of the same value of u_0. The proof ([2], p. 308) is based on a simple compactness argument: the set of all infinite R.F.'s is made into a compact, complete, metric space in which the two sequences $\{\omega_N^{(0)}\}$ and $\{\omega_N^{(1)}\}$ have the limit points $\omega^{(0)}$ and $\omega^{(1)}$ which are infinite R.F.'s with the desired properties.

References

1. R. L. Dobrushin, Description of a random field by means of conditional probabilities and the conditions governing its regularity, Theory of probability and its applications (English translation of Teoriia Veroiatn.), 13 (1968) 197–224.

2. ———, The problem of uniqueness of a Gibbs random field and the problem of phase transition, Functional Analysis and its applications (English translation of Funktionalny Analis), 2 (1968) 302–312.

3. David Ruelle, Statistical Mechanics, Benjamin, New York, 1969.

POSITIVE LIMIT SETS OF UNBOUNDED TRAJECTORIES

H. K. WILSON, Southern Illinois University at Edwardsville

1. Introduction. This paper is intended to be an expositional discussion of a topic in ordinary differential equations. In the introductory sections, the context of the discussion is established, and basic theorems are stated for the nonspecialist.

The differential equations under consideration have the form

$$x' = f(x, t), \tag{S}$$

where x is a real n-vector. In equation (S) and throughout the paper, primes denote differentiation with respect to t. We assume that f is defined, is continuous, and has continuous first partial derivatives with respect to its first n arguments on the interior of a cylindrical region $D=P\times(\alpha, +\infty)$. The set P is to be regarded as the union of a nonempty, open, connected set G and some subset of the boundary of G. The symbol α may stand for a real number or for the symbol $-\infty$.

If (x_0, t_0) in $\text{Int}(P)\times(\alpha, +\infty)$ is given, then the local existence and uniqueness theorem of ordinary differential equations states that there is solution $x=\phi_0(t)$ to the initial value problem

$$x' = f(x, t), \qquad x = x_0 \quad \text{when} \quad t = t_0. \tag{IVP}$$

The theorem guarantees that ϕ_0 is defined on some sufficiently short, nondegenerate interval $[t_0, t_1]$ and that ϕ_0 is the only solution of (IVP) defined on $[t_0, t_1]$. If the point $(\phi(t_1), t_1)$ is also in $\text{Int}(P)\times(\alpha, +\infty)$, then the initial value problem

$$x' = f(x, t), \qquad x = \phi_0(t_1) \quad \text{when} \quad t = t_1$$

will also have a solution $x=\phi_1(t)$ on some sufficiently short interval $[t_1, t_2]$. Consequently, the relations

$$x = \phi_0(t), \qquad t_0 \leqq t \leqq t_1,$$

and

$$x = \phi_1(t), \qquad t_1 \leqq t \leqq t_2$$

define another solution of the original problem (IVP) on the interval $[t_0, t_2]$. It is called a *forward extension* of the original solution $x=\phi_0(t)$. An extension to the left of t_0, rather than to the right of t_1, is called a *backward extension* of ϕ_0. One concludes, by applying transfinite induction to repeated forward and backward extensions, that there is a unique solution $x=\phi(t)$ to (IVP) which has a most inclusive interval of definition. This solution is called the *maximal solution* of (IVP). Let us denote the endpoints of the interval of definition for ϕ by a_ϕ and z_ϕ. Then (a_ϕ, z_ϕ) is contained in $(\alpha, +\infty)$ and ϕ maps this interval into $\text{Int}(P)$. It is a consequence of the extension method outlined above that the graph of ϕ

is not contained in any compact subset of $\text{Int}(P)\times(\alpha, +\infty)$. Thus, as $t\to z_\phi-$ or as $t\to a_\phi+$, either the point $(\phi(t), t)$ on the graph of ϕ approaches the boundary of D or else $|(\phi(t), t)|\to+\infty$. For this reason, z_ϕ and a_ϕ are referred to as *escape times*, in the future and in the past, respectively (Figure 1).

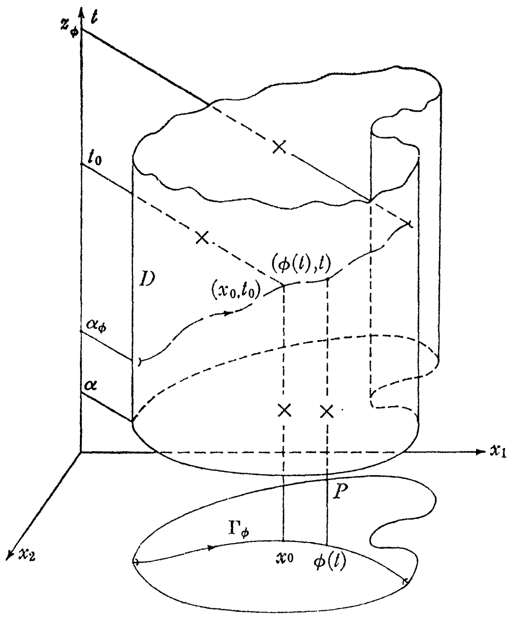


FIG. 1

In general, one must expect some of the solutions of (S) to have infinite future escape times and other solutions to have finite future escape times. Tests for the nonfiniteness of escape times are frequently based upon the fact that a solution $x=\phi(t)$ has $z_\phi=+\infty$ if $\phi(t)$ is contained in some compact subset of $\text{Int}(P)$ for $t_0\leqq t<z_\phi$. This result is intuitively evident if P is a subset of R^2. In this case, the point $(\phi(t), t)$ on the graph of ϕ cannot approach the boundary of D. Consequently, $(\phi(t), t)$ rises indefinitely with increasing t and $z_\phi=+\infty$ (Figure 2).

It can happen that a solution $x=\phi(t)$ has $z_\phi=+\infty$ even though $\phi(t)$ is not contained in any compact subset of $\text{Int}(P)$ for $t_0\leqq t<z_\phi$. The occurrence of this phenomenon can be detected by applying a generalization of the test described above: if, for each compact subinterval I of $[t_0, z_\phi)$, there is a compact subset $C(I)$ of $\text{Int}(P)$ such that $\phi(t)$ is in $C(I)$ for all t in I, then $z_\phi=+\infty$.

As a consequence of studying the system (S), one would hope, ideally, to compute *all* of its solutions and to describe all the associated graphs geometrically. For all but some of the simplest systems (S), however, this is a difficult goal to achieve. A less ambitious and more tractable goal is the qualitative description of the trajectories corresponding to solutions of (S). The *trajectory* Γ_ϕ

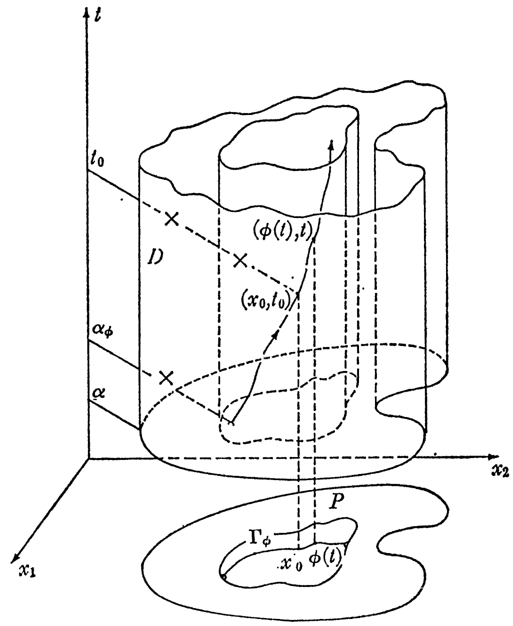


FIG. 2

corresponding to a solution $x=\phi(t)$ of (S) is the curve in P which ϕ represents; it is the projection of the graph of ϕ onto P. At least one trajectory of (S) passes through every point of P.

2. Limit Sets. The notion of limit set had its origin in the work of Poincaré [1] and had its general definition in G. D. Birkhoff's studies of dynamical systems [2]. This notion, which is central to topological discussions of ordinary differential equations, is a generalization of the limit concept for vector-valued functions of one real variable.

Suppose that $x=\phi(t)$, $a_\phi<t<z_\phi$, is a solution (S). If $\phi(t)\to c$ as $t\to z_\phi-$, then the singleton $\{c\}$ is called the positive limit set of Γ_ϕ. If $\phi(t)$ has no left hand limit at z_ϕ, then one considers sequences of times $t_k'<z_\phi$ such that $t_k'\to z_\phi-$ as $k\to+\infty$. The limit points of the sequence $\{t_k'\}$ are called *positive limit points* for Γ_ϕ. The *positive limit set* Ω_ϕ for Γ_ϕ is the set of all such points. Thus, one may say that c is a positive limit point for Γ_ϕ if there is a sequence of distinct times t_k such that $t_k\to z_\phi-$ and $\phi(t_k)\to c$ as $k\to+\infty$. An analogous *negative limit set* is defined for Γ_ϕ by considering $\phi(t)$ as $t\to a_\phi+$. Positive and negative limit sets are frequently called *ω-limit sets* and *α-limit sets*, respectively, in the literature of differential equations and topological dynamics.

As an illustration of the definition, consider the solution $x=(1+e^{-t})\sin t$, $y=(1+e^{-t})\cos t$ of the equations

$$x' = -x + y + \sin t, \qquad y' = -x - y + \cos t.$$

The unit circle in the xy-plane is the ω-limit set of the associated trajectory, and its α-limit set is empty.

What are the general properties of limit sets? It is convenient to answer this question for positive limit sets only. Let $x=\phi(t)$ be a solution of (S) and restrict it to a forward interval $[t_0, z_\phi)$. We denote the restricted function by ϕ^+ and its curve by Γ_ϕ^+. The limit set Ω_ϕ is studied by analyzing the forward half-trajectory Γ_ϕ^+. The choice of the point of restriction t_0 is immaterial.

At the outset one may say that Ω_ϕ is closed, although possibly empty. If $\Gamma_\phi{}^+$ lies in a bounded subset of P, then Ω_ϕ is compact and (in view of the Bolzano-Weierstrass Theorem) nonempty. The boundedness of Γ_ϕ^+ also implies that Ω_ϕ is connected. If Ω_ϕ were not connected, then it could be decomposed into two compact sets, separated by a positive distance. Then, since Γ_ϕ^+ is bounded and arcwise connected, it would be possible to construct a convergent sequence of points $\phi(t_k)$ with a limit not in Ω_ϕ. Using again a connectedness argument, one can show that Ω_ϕ is the smallest compact set which $\phi(t)$ approaches as $t\to z_\phi-$ provided, as it was above, that Γ_ϕ^+ is bounded.

The properties of Ω_ϕ when Γ_ϕ^+ is unbounded are discussed in Section 4 below.

A word of caution is in order when P is not all of R^n. There are two topologies which can reasonably be involved in statements about limit sets: the usual euclidean topology for R^n and the relative topology which it induces on P. Some statements can be interpreted equally well in either topology with entirely different meanings.

3. Autonomous systems. The system (S) is called autonomous if f is a function of x only. In this case we have $f(x, t)=g(x)$, say, and we replace the equation (S) by

$$x' = g(x). \tag{A}$$

Autonomous equations have special topological characteristics. The two most important of these for our purposes are that

(i) no trajectory of (A) can have a point in common with a different trajectory, and

(ii) a trajectory which intersects itself is either a point or a Jordan curve.

A point trajectory is called a *critical point*; it corresponds to a constant solution of (A). A simple, closed trajectory is called a *periodic orbit*; it corresponds to a periodic solution of (A).

We observed above, for solutions $x=\phi(t)$ of general equations (S), that if $\phi(t)\to c$ as $t\to z_\phi-$, then $\Omega_\phi=\{c\}$. More can be said for autonomous equations: if $x=\phi(t)$ is a solution of the autonomous equation (A) and if ϕ has a left hand limit c at z_ϕ, then c is a critical point for (A). This statement is not generally true for *non*autonomous systems. For example, the point (1, 0) is the only ω-limit point for the solution $x=e^{-t}+1$, $y=e^{-t}$ of the system

$$x' = -y, \qquad y' = -(2+e^t)y + x.$$

Nevertheless, the relations $x\equiv 1$, $y\equiv 0$ do not define a solution of the system.

Poincaré and Bendixson (see [1] and [3]) studied two-dimensional, autonomous equations (A). They showed that if the ω-limit set of a bounded trajectory in the plane contains no critical point, then this ω-limit set consists of precisely one periodic orbit.

Both critical points and periodic orbits possess a property that is shared by all limit sets for trajectories of autonomous systems: *invariance.* A set M in P is called *positively invariant* with respect to (A) if, for each point x_0 in M, the forward half-trajectory through x_0 is entirely contained in M. One defines the expression *negatively invariant* in a similar way, and the unqualified expression *invariant* means both positively and negatively invariant.

LaSalle [4] has devised a method for finding limit sets by studying invariant sets. Let us suppose, for example, that a half-trajectory Γ_ϕ^+ of (A) lies in a compact subset K of $\operatorname{Int}(P)$ and that K is positively invariant. LaSalle's theorem may be stated as follows:

If there exists a continuously differentiable, real-valued function V satisfying

$$\sum_{k=1}^{n} \frac{\partial V(x)}{\partial x_k} g_k(x) \leqq 0$$

for all x in K, then Ω_ϕ is contained in the union of all the invariant subsets of

$$Z = \left\{x : x \in K \text{ and } \sum_{k=1}^{n} \frac{\partial V(x)}{\partial x_k} g_k(x) = 0\right\}.$$

This theorem is used in the construction of Example 2, Section 5. The expression $\sum_{k=1}^{n} (\partial V(x)/\partial x_k) g_k(x)$ is called the *derivative of V along solutions of* (A).

4. The case of unbounded trajectories. Nemytskii and Stepanov [5] have shown by examples that the ω-limit set for an unbounded half-trajectory Γ_ϕ^+ of (A) need not be empty, and it can be disconnected. In each of their examples, however, the limit set Ω_ϕ lies in the boundary of P, that is, $\Omega_\phi \cap \operatorname{Int}(P) = \varnothing$. Now we commented above that statements about limit sets should be interpreted with regard to both the usual euclidean topology and the relative topology which it induces on P. Having in mind this observation and the fact that it is not unusual to study the differential equation (S) with P open in R^n, one may ask for examples of equations (S) with the following set of properties:

(i) There is a half-trajectory Γ_ϕ^+ which lies in no compact subset of P;
(ii) Γ_ϕ^+ *has* an ω-limit point in $\operatorname{Int}(P)$;
(iii) the limit set Ω_ϕ is connected.

At the same time, it is also reasonable to ask for examples of equations (S) with properties (i), (ii), and
(iii)′ the limit set Ω_ϕ is disconnected.

Such examples are easy to give in the nonautonomous case. The linear system

$$x' = -\frac{y}{t}, \qquad y' = tx + \frac{y}{t}$$

on $D=R^2\times(0, +\infty)$ has $x=\cos t$, $y=t\sin t$ for a solution. The corresponding ω-limit set consists of the points on the lines $x=\pm 1$. The trajectory of the solution

$$x = -1 + t\sin^2 t, \qquad y = t\sin t$$

of the system

$$x' = \frac{x}{t} + 2y\cos t + \frac{1}{t}, \qquad y' = \frac{y}{t} + t\cos t$$

on $D=R^2\times(0, +\infty)$ has, on the other hand, only the points of the line $x=-1$ for its ω-limit set.

It is more difficult to construct autonomous equations (A) which have solutions with the desired properties (i), (ii), (iii), and (i), (ii), (iii)$'$. Before proceeding to do so, let us summarize the general properties of Ω_ϕ when Γ_ϕ^+ is not contained in any compact subset of P.

The limit set Ω_ϕ is certainly closed. If $\Omega_\phi \cap \mathrm{Int}(P)\neq\emptyset$, then $z_\phi=+\infty$ [6, Theorem 12.1]. Thus the mere existence of an interior ω-limit point for Γ_ϕ eliminates the possibility of a finite escape time for ϕ. Suppose next, for contradiction, that Ω_ϕ is contained in a compact subset of $\mathrm{Int}(P)$. Then there exist compact sets B and C such that

$$\Omega_\phi \subset \mathrm{Int}(B) \subset B \subset \mathrm{Int}(C) \subset C \subset \mathrm{Int}(P).$$

Since $\Gamma_\phi^+ \not\subset C$ by assumption and $\Gamma_\phi^+ \cap \mathrm{Int}(B)\neq\emptyset$ by definition of Ω_ϕ, there exists a sequence $\{t_n\}$ such that $\phi(t_n)$ is in $C-\mathrm{Int}(B)$ and $t_n\to+\infty$ as $n\to+\infty$. Thus

$$\Omega_\phi \cap \overline{(R^n - B)} \neq \emptyset \quad \text{and} \quad \Omega_\phi \not\subset \mathrm{Int}(B).$$

This is a contradiction. It follows that Ω_ϕ is contained in no compact subset of P. The examples in the next section show that Ω_ϕ may be connected or it may be disconnected with its components separated by a positive distance.

5. Construction of examples. We begin by displaying an autonomous equation (A), with $P=R^2-\{(0, 1)\}$, which has properties (i), (ii), and (iii).

Example 1. By using polar coordinates, one can verify that the unit circle in the uv-plane is the ω-limit set of every trajectory of the equation

$$\zeta' = (-1+i)\zeta + \frac{\zeta}{|\zeta|}, \qquad \zeta = u + iv \neq 0. \tag{1}$$

The substitution of

$$\zeta = \frac{i - z}{i + z}, \qquad z = x + iy, \tag{2}$$

into (2) gives rise to the equation

$$z' = -\frac{1}{2}(1+i)(i-z)(i+z) + \frac{i(i-z)(i+z)\,|\,i+z\,|}{2\,|\,i-z\,|}, \qquad z \neq i. \tag{3}$$

Since the inverse transformation

$$z = i\,\frac{1-\zeta}{1+\zeta}$$

of (2) maps the uv-plane conformally onto the xy-plane in such a way that the unit circle maps onto the x-axis, it follows that the trajectory for each unbounded solution

$$z(t) = i\,\frac{1 - e^{it}(1+ce^{-t})}{1 + e^{it}(1+ce^{-t})}, \qquad c \neq 0, \tag{4}$$

of (3) has the x-axis for its ω-limit set. The x-axis is traced by the solution

$$z(t) = i\,\frac{1-e^{it}}{1+e^{it}} = \tan\frac{t}{2}, \qquad -\frac{\pi}{2} < t < \frac{\pi}{2}, \tag{5}$$

and the corresponding ω-limit set is empty. Note, incidentally, that the solutions (4) have infinite escape times, but the solution (5) has a finite escape time.

The next example is an autonomous equation (A), with $P = R^2 - \{(1, 0)\}$, which has properties (i), (ii), and (iii)$'$.

Example 2. Define $V_0(u, v) \equiv u^2 + v^2 - 4$ and $V_I(u, v) \equiv u^2 + (v-1)^2 - 1$. Let C_0 and C_I denote the circles defined by $V_0(u, v) = 0$ and $V_I(u, v) = 0$, respectively, and let G denote the open region between them (Figure 3).

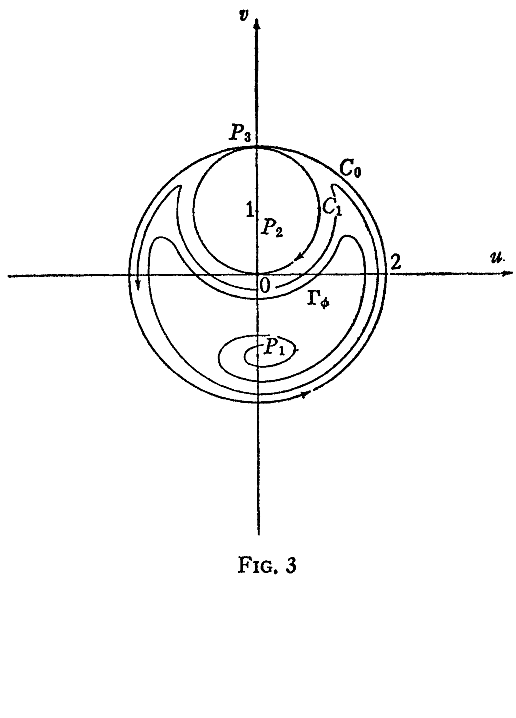


FIG. 3

We first establish that every trajectory of the system

$$\begin{aligned} u' &= 2uvV_IV_0 - 2vV_I - (2v-2)V_0, \\ v' &= -2u^2V_IV_0 + 2uV_I + 2uV_0 \end{aligned} \tag{6}$$

which initiates in G has $C_0 \cup C_I$ for its ω-limit set.

One easily calculates that the derivatives of V_0 and V_I along solution of the system (6) are given by

$$V_0' = 4uV_0 \tag{7}$$

and

$$V_I' = 4uV_I(uV_0 - 1). \tag{8}$$

Thus C_0 and C_I are invariant sets for (6). If M is an invariant set for (6) which is contained in the line $u=0$, then $u'=0$ for (u, v) in M. Thus M can only contain one or more of the critical points P_1, P_2, P_3 with respective ordinates

$$v_1 = -\frac{1}{4} - \sqrt{1+\frac{1}{16}}, \qquad v_2 = -\frac{1}{4} + \sqrt{1+\frac{1}{16}}, \qquad v_3 = 2.$$

The union of the invariant sets of (6) which are contained in

$$Z = \{(u, v) : (u, v) \in \overline{G} \quad \text{and} \quad (V_IV_0)' \equiv 4V_IV_0^2u^2 = 0\}$$

is $C_I \cup C_0 \cup \{p_1\}$. Since $V_IV_0 < 0$ on G and $(V_IV_0)' > 0$ on $G - Z$, it follows from LaSalle's theorem [4] that $C_I \cup C_0$ contains the ω-limit set of every nonconstant trajectory of (6) in G, and $\{p_1\}$ is the negative ($\alpha-$) limit set of every such trajectory. Suppose now, for contradiction, that there is a nonconstant solution $\psi = (u, v)$ of (6) with $\psi(t)$ in G such that $\Omega_\psi \neq C_I \cup C_0$. Then $\lim_{t\to+\infty} \psi(t) = p_3$. There are two possibilities: either $u(t) > 0$ for all sufficiently large t or $u(t) < 0$ for all such t. Consider the latter possibility. Then there exists a $T \geqq 0$ such that $t > T$ implies $-1 < u(t)\,V_0(u(t), v(t)) - 1 < -\frac{1}{2}$. By inequality (8),

$$V_I(u(t), v(t)) \geqq V_I(u(T), v(T)) \exp \int_T^t 2\,|\,u(s)\,|\,ds,$$

and $V_I(u(t), v(t))$ does not approach zero as $t \to +\infty$. If $u(t) > 0$ for all sufficiently large t, a similar argument, using (7), produces a contradiction. Thus $C_I \cup C_0$ is the ω-limit set of every nonconstant trajectory of (6) in G.

The mapping

$$\zeta = 2i\frac{z+1}{z-1}, \qquad z = x + iy, \qquad \zeta = u + iv \tag{9}$$

or

$$u = \frac{4y}{(x-1)^2 + y^2}, \qquad v = 2\frac{x^2 + y^2 - 1}{(x-1)^2 + y^2}$$

is a conformal correspondence between the compactified uv- and xy-planes. The lines $x=-1$ and $x=0$ map onto the circles C_I and C_0, respectively, and the strip $-1<x<0$ maps onto the region G. If $\psi=(u, v)$ is a nonconstant solution of (6) with $\psi(t)$ in G, then $\phi=(x, y)$ defined by

$$x = \frac{u^2+v^2-4}{u^2+(v-2)^2}, \qquad y = \frac{4u}{u^2+(v-2)^2}$$

is an unbounded solution of the system

$$x' = f(x, y), \qquad y' = g(x, y) \tag{10}$$

obtained from (6) by means of the substitution (9), and Ω_ϕ consists of the lines $x=-1$ and $x=0$.

Systems with properties (i), (ii), and (iii) or (i), (ii) and (iii)$'$ and $n=3$ can be trivially constructed by appending the equation $z'=0$ to each of the systems (1) and (10). The next example makes use of a less trivial construction.

Example 3. The system

$$x' = -(y+x)(x^2+y^2), \qquad y' = -(y-x)(x^2+y^2) \tag{11}$$

has the form

$$\rho' = -2\rho^2, \qquad \theta' = \rho$$

in polar coordinates $\rho=x^2+y^2$, $\tan\theta=y/x$. Its nonzero solutions are given by

$$\rho = \frac{\rho_0}{1+2\rho_0 t}, \qquad \theta = \theta_0 + \ln\sqrt{1+2\rho_0 t}, \qquad \rho_0 \neq 0,$$

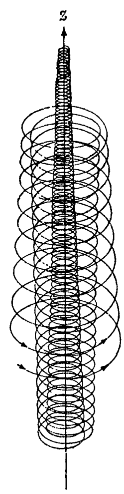


FIG. 4

and the corresponding trajectories are spirals in the xy-plane. If the equation

$$z' = x - y \tag{12}$$

is appended to the system (11), then

$$z = z_0 + \int_0^t \sqrt{\frac{2\rho_0}{1 + 2\rho_0 s}} \cos\left(\theta_0 + \frac{\pi}{4} + \ln \sqrt{1 + 2\rho_0 s}\right) ds.$$

The z-axis is a line of critical points for the system (11; 12). It is the ω-limit set of every nonconstant trajectory since $\rho \to 0$, $\theta \to +\infty$, and z oscillates unboundedly as $t \to +\infty$ (Figure 4).

References

1. H. Poincaré, Sur les courbes définies par une équation différentielle, Oeuvres, Vol. I, Paris, 1892, 3–84.
2. G. D. Birkhoff, Dynamical Systems, Colloquium Publications No. 9, American Math. Soc., 1927.
3. I. Bendixson, Sur les courbes définies par les équations différentielles, Acta Math., 24 (1901) 1–88.
4. J. P. LaSalle, Some extensions of Liapunov's second method, IRE Trans. on Circuit Theory, CT-7(1960) 520–527.
5. V. V. Nemytskii and V. V. Stepanov, Qualitative Theory of Differential Equations, Princeton University Press, 1960, 340–344.
6. T. Yoshizawa, Stability Theory by Liapunov's Second Method, Mathematical Society of Japan, Tokyo, 1966.

INTEGRAL TESTS FOR INFINITE SERIES

O. E. STANAITIS, St. Olaf College

1. Introduction. One of the most powerful tests for convergence of infinite series is the familiar Maclaurin-Cauchy integral test.

If $f(x)$ is a positive, continuous, and monotonically decreasing function defined for $x \geqq 1$, then

$$\sum_{x=1}^{\infty} f(x) \quad \text{and} \quad \int_1^{\infty} f(x)dx$$

converge and diverge together. In case of divergence the difference

$$\int_1^n f(x)dx - \sum_{x=1}^{n} f(x)$$

tends to a definite limit d $(0 < d < f(1))$ as n tends to infinity.

This relationship plays an important role in approximating convergent and divergent series, but its scope of applications is very limited.

a. Bromwich's Theorem. A first interesting extension of the test was given by F. J. I'A. Bromwich [1].

If

(i) $f(x)$ is positive and tends steadily to zero;
(ii) $g(x)$ is positive and tends steadily to infinity;
(iii) $g'(x)$ tends steadily to zero;
(iv) $\int_1^\infty f(x)g'(x)dx$ is convergent, then

$$\int_1^n F(x)dx - \sum_{x=1}^n F(x),$$

where $F(x) = f(x)e^{ig(x)}$, tends to a finite limit as n tends to infinity. Here $i^2 = -1$, and the familiar relation $e^{iv} = \cos v + i \sin v$ is used in the following examples.

b. G. H. Hardy's Generalization and Examples. Hardy generalized Bromwich's theorem in the following way [2]:

If

(i) $F(x)$ possesses a continuous derivative $F'(x)$,
(ii) $F(x) \to 0$ as $x \to \infty$,
(iii) $\int_1^\infty |F'(x)|dx$ is convergent, then

$$\int_1^n F(x)dx - \sum_{x=1}^n F(x)$$

tends to a definite limit as $n \to \infty$.

That Hardy's theorem represents a more general test is indicated by the example

$$F(x) = x^{-a}\exp[ix^b \cos(\log x)], \text{ where } 0 < b < 1,$$

where Bromwich's test does not apply.

If the functions in Bromwich's theorem are such that the integrals

$$\int_1^\infty f'(x)dx, \qquad \int_1^\infty f(x)g'(x)dx$$

are convergent, then from

$$F'(x) = [f'(x) + if(x)g'(x)]e^{ig(x)}$$

it follows immediately that $\int_1^\infty |F'(x)|dx$ in Hardy's theorem is convergent.

Both theorems apply to many interesting series such as

$$\sum_{n=1}^\infty \frac{(\log n)^p}{n^{1+ai}}$$

where a is real.

But neither theorem is sufficiently general to deal with many other simple series, for example,

$$\sum_{n=1}^{\infty} \frac{e^{in^a}}{n^b} \qquad 0 < a < 1, \quad b > 0.$$

The above series converges if $a+b>1$, but the theorems require an additional condition $b>a$.

c. HARDY'S SECOND THEOREM. Hardy proved that if for some integer $s\geqq 1$ the integral $\int_1^\infty |F^{(s)}(x)|dx$ is convergent, then

$$\int_1^n F(x)dx - \sum_{x=1}^{n} F(x)$$

tends to a finite limit as $n\to\infty$.

It is easily seen that for

$$F(x) = x^{-b}e^{ix^a}, \qquad 0 < a < 1, \quad b > 0$$

it follows that

$$F^{(s)}(x) \sim (ia)^s x^{s(a-1)} - be^{ix^a}.$$

If s is chosen such that $s(a-1)-b<-1$ $[s>(1-b)/(1-a)]$ the integral $\int_1^\infty |F^{(s)}(x)|dx$ converges. Consequently

$$\int_1^\infty \frac{e^{ix^a}}{x^b} dx \quad \text{and} \quad \sum_{x=1}^{\infty} \frac{e^{ix^a}}{x^b}$$

both converge if $a+b>1$.

The generalized test, however, does not apply to the familiar convergent series for $a=1$, nor does it give any information when $a>1$.

The object of this note is to point out more general tests. We start with a more general theorem.

2. **Theorem.** *Let $f(x)$ and its derivatives $f^{(p)}(x)$, defined for $x\geqq 1$, be continuous and tend to zero as $x\to\infty$. Let $P_{2s}(x)$ and $P_{2s+1}(x)$ denote the series*

$$P_{2s}(x) = (-1)^{s-1}\sum_{n=1}^{\infty} \frac{2\cos(2n\pi x)}{(2n\pi)^{2s}},$$

$$P_{2s+1}(x) = (-1)^{s-1}\sum_{n=1}^{\infty} \frac{2\sin(2n\pi x)}{(2n\pi)^{2s+1}}.$$

Then the difference $\int_1^n f(x)dx - \sum_{x=1}^n f(x)$ tends to a definite limit as $n\to\infty$ if and only if the integral $\int_1^\infty P_k(x)f^{(k)}(x)dx$ converges, where $k=2s$ or $2s+1$.

Proof. From Euler's summation formula

$$\sum_{x=1}^{n} f(x) = \int_1^n f(x)dx + \frac{1}{2}[f(n) + f(1)] + \frac{B_2}{2!}[f'(n) - f'(1)]$$
$$+ \frac{B_4}{4!}[f'''(n) - f'''(1)] + \cdots + \frac{B_{2s}}{(2s)!}[f^{(2s-1)}(n) - f^{(2s-1)}(1)] + R_k,$$

where B_{2r}, $r = 1, 2, 3, \cdots$ are Bernoulli numbers and

$$R_k = \int_1^n P_{2s+1}(x)f^{(2s+1)}x\, dx = -\int_1^n P_{2s}(x)f^{(2s)}(x)\, dx,$$

it follows that

$$\lim_{n\to\infty}\left(\int_1^n f(x)\, dx - \sum_{x=1}^{n} f(x)\right) = \frac{-1}{2}f(1) + \frac{B_2}{2!}f'(1) + \frac{B_4}{4!}f^{(3)}(1) + \cdots$$
$$+ \frac{B_{2s}}{(2s)!}f^{(2s-1)}(1) + \int_1^\infty P_k(x)f^{(k)}(x)\, dx.$$

Since by hypothesis the derivatives $f^{(p)}(1)$ exist, it follows that the limit exists if the integral on the right converges. And conversely, the existence of the limit implies the convergence of the integral. This proves the theorem.

3. Corollaries. It is apparent that the infinite series converges if both integrals of Theorem 2 exist, and diverges if one of them does not exist. In particular, the integral

$$\int_1^\infty P_k(x)f^{(k)}(x)\, dx$$

converges if the integral $\int_1^\infty |f^{(k)}(x)|\, dx$ exists. This leads to three corollaries.

COROLLARY 1. *The series* $\sum_{x=1}^\infty f(x)$ *converges if the integrals*

$$\int_1^\infty f(x)\, dx \quad and \quad \int_1^\infty |f^{(k)}(x)|\, dx$$

exist $(k \geqq 1)$.

COROLLARY 2. *The series* $\sum_{x=1}^\infty f(x)$ *converges if* $\int_1^\infty f(x)\, dx$ *exists and*

$$\int_1^\infty f^{(p)}(x)\sin(2k\pi x)dx = O(1/k^r)$$

for $r > 0$ *and* $p \geqq 1$.

Proof. The statement follows immediately from the relation

$$\int_1^\infty P_p(x)f^{(p)}(x)\, dx = (-1)^{s-1}\sum_{k=1}^\infty \frac{2}{(2k\pi)^p}\int_1^\infty f^{(p)}(x)\sin(2k\pi x)dx,$$

where $p=2s+1$.

COROLLARY 3. *If the integral* $\int_1^\infty f(x)dx$ *converges and*

$$\lim_{k\to\infty}\int_1^\infty f'(x)\sin(2k\pi x)dx = A \neq 0$$

exists, then the series $\sum_{x=1}^\infty f(x)$ *diverges.*

Proof. From the relation $\int_1^\infty P_1(x)f'(x)\,dx = -\sum_{k=1}^\infty 1/k\int_1^\infty f'(x)\sin(2k\pi x)\,dx$ it follows that the series diverges if

$$A-\epsilon < \int_1^\infty f'(x)\sin(2k\pi x)dx < A+\epsilon$$

holds for $k>k_0$.

NOTE 1. Corollary 1 contains both Hardy's theorems. The test applies to infinite series that often occur in advanced calculus and certainly is preferable to Cauchy's convergence criterion which usually requires ingenuity and a lot of calculation [3].

NOTE 2. A proof that summation and integration may be interchanged in proofs of corollaries 2 and 3 has been omitted. It is sometimes overlooked that, in general, uniform convergence of a series $\sum_{n=1}^\infty f_n(x)$ does not assure the equality

$$\int_a^\infty g(x)\left[\sum_{n=1}^\infty f_n(x)\right]dx = \sum_{n=1}^\infty\int_a^\infty g(x)f_n(x)dx$$

even if both sides converge. In our case a sufficient condition (existence of the integral $\int_a^\infty |g(x)|dx$) is not satisfied.

NOTE 3. If in Corollary 3 $\lim_{k\to\infty}\int_1^\infty f'(x)\sin(2k\pi x)dx$ does not exist, the test is inconclusive. The series might still converge if, for example, the integrand of $\int_1^\infty f'(x)\sin(2k\pi x)dx$ is periodic and such that groups of terms have opposite signs. A simple case in point would be if

$$\int_1^\infty f'(x)\sin(2k\pi x)dx = (-1)^k(A+\epsilon_k),$$

where $\lim_{k\to\infty}\epsilon_k=0$.

4. Example. Let us apply Corollary 2 to Hardy's example

$$f(x) = x^{-b}e^{ix^a}$$

when $0<b\leqq 1$ and $0<a<2$.

It is easily checked that the first integral $\int_1^\infty f(x)dx$ exists if $a+b>1$. The second integral on which convergence of the series $\sum_{x=1}^\infty f(x)$ depends can be put in the form

$$\begin{aligned}\int_1^\infty f'(x)\sin(2k\pi x)dx \\ (1)\qquad\qquad = ia\int_1^\infty \frac{e^{ix^a}\sin(2k\pi x)}{x^{b-a+1}}dx - b\int_1^\infty \frac{e^{ix^a}\sin(2k\pi x)}{x^{b+1}}dx.\end{aligned}$$

If $0 < a \leqq 1$ it is easily seen that the first integral on the right exists and the second is absolutely convergent. Integration by parts yields

$$\begin{aligned}\int_1^\infty \frac{e^{ix^a}\sin(2k\pi x)}{x^{b-a+1}}dx \\ (2)\qquad = \frac{1}{2k\pi}\left(e^i + ia\int_1^\infty \frac{e^{ix^a}\cos(2k\pi x)}{x^{b-2a+2}}dx - (b-a+1)\int_1^\infty \frac{e^{ix^a}\cos(2k\pi x)}{x^{b-a+2}}dx\right).\end{aligned}$$

Hence

$$\int_1^\infty f'(x)\sin(2k\pi x)dx = O(1/k)$$

and it follows that the series $\sum_{x=1}^\infty f(x)$ converges if $a+b>1$ and $0<a\leqq 1$.

The case $1<a<2$ is somewhat more difficult. If we assume that $b-a+1>0$, then (2) holds also for $1<a<2$ and the second integral on the right is absolutely convergent. To show that the integral on the left of (2) exists we make a change of variable $x^a = t$:

$$\begin{aligned}\int_1^\infty \frac{e^{ix^a}\sin(2k\pi x)}{x^{b-a+1}}dx &= \frac{1}{a}\int_1^\infty \frac{\sin(2k\pi t^{1/a})}{t^{b/a}}e^{it}dt \\ &= \frac{-1}{ai}\int_1^\infty e^{it}\left(\frac{\sin(2k\pi t^{1/a})}{t^{b/a}}\right)' dt.\end{aligned}$$

By repeated integration by parts the existence of the integral is easily established. It follows as a corollary from (2) that the first integral on the right also exists.

Furthermore, from the existence of the integral

$$\int_1^\infty \frac{e^{ix^a}}{x^{b-2a+2}}dx$$

and the familiar relation

$$\lim_{k\to\infty}\int_1^n \frac{e^{ix^a}}{x^{b-2a+1}}\cos(2k\pi x)dx = 0$$

which holds for all finite n, it follows that the first integral on the right of (2) is bounded for all k. Consequently

$$\int_1^{\infty} f'(x) \sin(2k\pi x)dx = O(1/k)$$

which completes the proof.

Thus for $1<a<2$ the series $\sum_{x=1}^{\infty} f(x)$ converges if $b-a+1>0$ ($a+b>1$ is satisfied). Note that convergence in case $a=1$ does not depend on any conditions, since $a+b>1$ and $b-a+1>0$ are always satisfied.

The example shows that the scope of applications of Corollary 2 is considerably wider than that of Corollary 1.

References

1. T. J. I'A. Bromwich, Proc. Lond. Math. Soc., (2) 6 (1908) 327.
2. G. H. Hardy, Proc. Lond. Math. Soc., (2) 9 (1910) 126.
3. Angus E. Taylor, Advanced Calculus, Ginn 1955, p. 585 (Prob. 17 and 18).

MATHEMATICAL NOTES

EDITED BY DAVID DRASIN

Manuscripts for this Department should be sent to David Drasin, Division of Mathematical Sciences, Purdue University, Lafayette, IN 47907.

THERE IS AN ELEMENTARY PROOF OF PEANO'S EXISTENCE THEOREM

WOLFGANG WALTER, Mathematisches Institut der Universität, Karlsruhe, Germany

A "Research Problem" by H. C. Kennedy (see [2]) stimulated the author to publish the following proof, which he discovered more than 10 years ago and since has used in lectures. We shall first present the proof, which is constructive and deserves, in our opinion, the adjective "elementary" without restriction, because the construction of a monotone decreasing sequence of approximate solutions avoids equicontinuous families and an appeal to the Ascoli lemma. I am hesitant in making the same statement on Peano's or Perron's proof. Critical and historical remarks will be given at the end of the paper.

PEANO'S EXISTENCE THEOREM. *Let J be the interval $0 \leqq t \leqq T$ $(T>0)$ and let $f(t, x): J\times R \to R$ be a continuous and bounded function. Then there exists, for a given $u_0 \in R$, at least one continuously differentiable function $u(t): J\to R$ satisfying*

$$u' = f(t, u) \quad \textit{in} \quad J \quad \textit{and} \quad u(0) = u_0. \tag{1}$$

The proof utilizes a variant of the Euler-Cauchy polygon method. Let $h=T/n>0$ be given and $t_i = ih$ $(i=0, 1, \cdots, n)$. In the polygon method one constructs a sequence $(v_i)_0^n$ according to

$$v_0 = u_0, \qquad v_{i+1} = v_i + hf(t_i, v_i) \quad (i = 0, 1, \cdots, n-1).$$

We use instead the formula

(2) $v_0 = u_0, \quad v_{i+1} = v_i + h \max\{f(t, x): t_i \leqq t \leqq t_{i+1}, v_i - 3Mh \leqq x \leqq v_i + Mh\},$

where $|f(t, x)| \leqq M$ in $J \times R$; as in the classical Euler-Cauchy polygon method the approximate solution $v(t)$ is constructed by joining the points (t_i, v_i) $(i = 0, 1, \cdots, n)$ by a polygonal line. This construction, carried out for the parameters $h, h/2, h/4, \cdots$, leads to a monotone decreasing sequence of approximate solutions.

In proving the last statement we use the following notation: $h = T/n > 0$ is fixed, $t_i = ih$ $(i = 0, \frac{1}{2}, 1, \frac{3}{2}, \cdots, n)$, $(v_i) = (v_0, v_1, \cdots, v_n)$ is constructed according to (2) with respect to the parameter h, $(w_i) = (w_0, w_{1/2}, w_1, w_{3/2}, \cdots, w_n)$ is constructed similarly, but with respect to the parameter $h/2$, $v(t)$ and $w(t)$ are continuous, piecewise linear functions satisfying $v(t_i) = v_i$ for $i = 0, 1, 2, \cdots, n$ and $w(t_i) = w_i$ for $i = 0, \frac{1}{2}, 1, \cdots, n$. Let $R(t, x; h)$ be the rectangle $[t, t+h] \times [x - 3Mh, x + Mh]$ and let $R_i = R(t_i, v_i; h)$ $(i = 0, 1, \cdots, n)$, $S_i = R(t_i, w_i; h/2)$ $(i = 0, \frac{1}{2}, 1, \cdots, n)$. In this notation the sequences (v_i) and (w_i) are defined as follows:

$$v_0 = w_0 = u_0, \qquad v_{i+1} = v_i + h \max f(R_i), \qquad w_{i+1/2} = w_i + (h/2) \max f(S_i).$$

Here the standard notation $\max f(A) = \max\{f(t, x): (t, x) \in A\}$ was used.

We shall prove, by induction on i, that $w(t) \leqq v(t)$ in J. Let us assume that i is a nonnegative integer and that $w(t) \leqq v(t)$ for $0 \leqq t \leqq t_i$. Then there are two cases to be considered,

$$\text{(a) } w_i \leqq v_i - Mh; \qquad \text{(b) } v_i - Mh < w_i \leqq v_i.$$

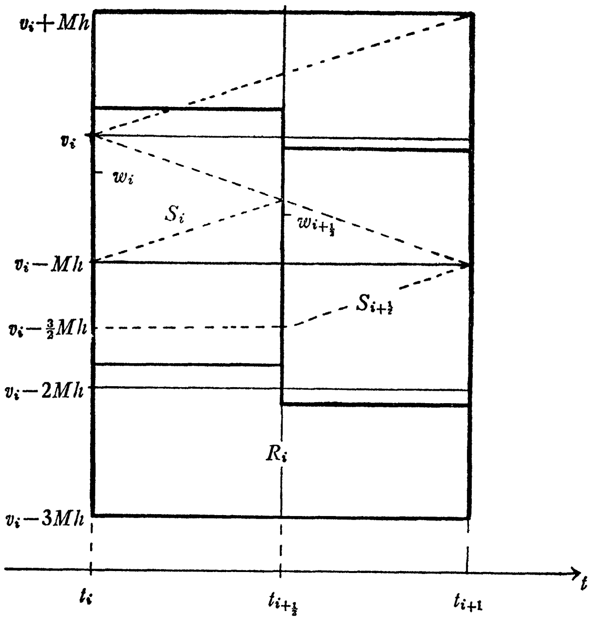
In case (a) we have $w(t) \leqq v(t)$ for $t_i \leqq t \leqq t_i + h/2$, since $|v'|, |w'| \leqq M$. In case (b) we have $S_i \subset R_i$ and hence $w' \leqq v'$ for $t_i < t < t_i + h/2$; again the inequality $w \leqq v$ in $[t_i, t_{i+1/2}]$ follows. In essentially the same way one shows that $w \leqq v$ in $[t_{i+1/2}, t_{i+1}]$. Since $v(t_{i+1/2}) \leqq v_i + Mh/2$ there are two cases

$$\text{(a) } w_{i+1/2} \leqq v_i - \tfrac{3}{2}Mh; \qquad \text{(b) } v_i - \tfrac{3}{2}MH < w_{i+1/2} \leqq v_i + \tfrac{1}{2}Mh.$$

In case (a) it follows as above that $w(t) \leqq v(t)$ for $t_{i+1/2} \leqq t \leqq t_{i+1}$, while in case (b) we have $S_{i+1/2} \subset R_i$ and therefore $w'(t) \leqq v'(t)$ for $t_{i+1/2} < t < t_{i+1}$. In the figure case (a) is indicated by the punctuated lines, while the two rectangles S_i, $S_{i+1/2}$ correspond to case (b). So far we have proved that the inequality $w(t) \leqq v(t)$ holds in $[0, t_{i+1}]$; it follows then by induction that this inequality is true in J.

The rest of the proof is a matter of routine. Let $h_k = T \cdot 2^{-k}$, $k = 1, 2, \cdots$, and let $v_k(t)$ be the piecewise linear, continuous function constructed according to (2) with respect to $h = h_k$. We have proved that $v_{k+1}(t) \leqq v_k(t)$ in J. Furthermore, v_k satisfies a Lipschitz condition with a Lipschitz constant M independent of k, and $v_k(t)$ is bounded below by $u_0 - Mt$. Therefore $u(t) = \lim_{k \to \infty} v_k(t)$ exists, the convergence being uniform in J. The function u is continuous (even Lipschitz continuous) in J; it is the desired solution of the initial value problem, as it will be shown now.

With the exception of a finite number of points, $v_k'(t)$ exists and, according to (2),



$$v_k'(t) = f(t', x'), \quad \text{where} \quad |t - t'| \leqq h_k, \quad |v_k(t) - x'| \leqq 3Mh_k.$$

If $d(s)$ denotes a modulus of continuity for f,

$$|f(t, x) - f(t', x')| \leqq d(|t - t'| + |x - x'|) \quad \text{in } J \times [u_0 - MT, u_0 + MT],$$

then $v_k'(t) = f(t, v_k(t)) + \alpha_k(t)$, where $|\alpha_k(t)| \leqq d(h_k + 3Mh_k)$ and hence

$$(3) \qquad v_k(t) = u_0 + \int_0^t f(s, v_k(s))ds + \beta_k(s), \quad |\beta_k(s)| \leqq d(h_k + 3Mh_k)T.$$

It follows immediately from (3) for $k \to \infty$ that

$$u(t) = u_0 + \int_0^t f(s, u(s))ds \qquad \text{in } J,$$

i.e., that u is a solution of the initial value problem (1).

REMARKS. (a) It is an easy exercise to prove the following statement: If $\bar{u}(t)$ is another solution of (1), then $\bar{u} \leqq v_k$ in J. Hence the solution u constructed above is indeed the maximal solution.

(b) It is not the aim of this paper to investigate the existence proofs by Peano (1886 and 1890) and Perron (1915). Nevertheless, the author is not in agreement with several critical remarks in [2] concerning these proofs. Perron's proof is correct. Furthermore the remark in [2] that Peano's second proof (1890) is based on successive approximation is incorrect.

(c) The theoretical basis for the proof given in this paper as well as for

Peano's first proof (1886) and Perron's proof is a theorem on differential inequalities which in the simplest case reads as follows (see, e.g., [3; p. 57]):

(A) *Let* $v(t)$, $w(t)$ *be differentiable in* J *and* $v(0) < w(0)$, $v'(t) - f(t, v) < w'(t) - f(t, w)$ *in* J. *Then* $v < w$ *in* J.

Due to this theorem the operator $L\varphi = (\varphi' - f(t, \varphi), \varphi(0) - u_0)$ is "monotone." Using an obvious interpretation of inequalities, Theorem (A) may simply be stated as "$Lv < Lw$ implies $v < w$." Therefore the inequalities $Lv < 0$, $Lw > 0$ characterize a subfunction v and a superfunction w such that $v < u < w$ for each solution u of (1).

The existence proofs mentioned above in connection with (A) cannot be transformed to systems of ordinary differential equations; the vector analog of (A) is not true in general. It is only true if the function

$$f(t, x) = (f_1(t, x_1, \cdots, x_n), \cdots, f_n(t, x_1, \cdots, x_n))$$

is "quasimonotone increasing in x", i.e., if $f_i(t, x_1, \cdots, x_n)$ is increasing in x_j for $i \neq j$; see, e.g., [3; p. 83]. We remark that the existence proof given here and Perron's proof carry over to systems whose right hand side $f(t, x)$ is quasimonotone increasing in x.

(d) In 1959, when the author carried out his "Habilitation" at the University of Karlsruhe, Germany, the "*Probevorlesung*" was still part of this academic procedure. In his *Probevorlesung*, entitled *Der Existenzsatz von Peano*, the author first gave the existence proof described in this paper. Independently and about the same time H. Grunsky found another constructive existence proof [1] which, though being different in the technical details, has basic ideas in common with our proof.

(e) Naturally, the minimum solution of (1) may be constructed in essentially the same way. In (2) one has to replace the maximum by the minimum, and the rectangle is now given by $[t_i, t_{i+1}] \times [v_i - Mh, v_i + 3Mh]$. If $(\bar{v}_k(t))$ is the sequence constructed in this way for $h_k = 2^{-k}T$, then $(\bar{v}_k)$ turns out to be a monotone increasing sequence, and $\lim \bar{v}_k(t)$ is the minimum solution of (1). Therefore, if u is any solution of (1), then

$$\bar{v}_k \leqq \bar{v}_{k+1} \leqq u \leqq v_{k+1} \leqq v_k \qquad \text{in } J$$

for all k. In other words, our method is a numerical procedure which yields monotone sequences of upper and lower bounds.

References

1. H. Grunsky, Ein konstruktiver Beweis für die Lösbarkeit der Differentialgleichung $y' = f(x, y)$ bei stetigem $f(x, y)$, Jahresber. Deutsch. Math. Ver., Abt. 1, 63 (1960) 78–84.

2. H. C. Kennedy, Is there an elementary proof of Peano's existence theorem for first order differential equations? this MONTHLY, 76 (1969) 1043–1045.

3. W. Walter, Differential- und Integral-Ungleichungen. Springer Tracts in Natural Philosophy, Vol. 2. Springer Verlag 1964.

THE BIDISK IS NOT BIHOLOMORPHIC TO THE BALL

J. J. HIRSCHFELDER, University of Washington, Seattle

The *Riemann Mapping Theorem* states that each open, bounded, simply connected subset of the complex plane $\boldsymbol{C}$ can be mapped biholomorphically (conformally) onto the open unit disk $\boldsymbol{D}=\{z\in\boldsymbol{C} \mid |z|<1\}$. The student may wonder whether a similar result holds in several variables. It does not; in fact, the bidisk $\boldsymbol{D}\times\boldsymbol{D}\subset\boldsymbol{C}^2$ and the unit ball $\boldsymbol{B}=\{(z, w)\in\boldsymbol{C}^2 \mid |z|^2+|w|^2<1\}$ are not biholomorphically equivalent. The classical proof may be found in [**1**, p. 96].

We give here a proof, based on a notion of Kobayashi [**2**], which uses only the following two facts from the elementary theory of one complex variable:

SCHWARZ LEMMA: *If $f:\boldsymbol{D}\rightarrow\boldsymbol{D}$ is holomorphic and $f(0)=0$, then $|f(z)|\leqq|z|$ for all $z\in\boldsymbol{D}$.*

HOMOGENEITY OF THE DISK: *If z and w are any two points of $\boldsymbol{D}$, then there is a biholomorphic map $f:\boldsymbol{D}\rightarrow\boldsymbol{D}$ such that $f(z)=w$.*

Of course, we are going to prove a theorem about holomorphic functions of several complex variables, but the reader need not know what such things are. However, he is asked to believe that they are differentiable, and that a composition of holomorphic functions is holomorphic.

Let $\boldsymbol{M}$ be an open set in $\boldsymbol{C}^n$. We define a function $k_{\boldsymbol{M}}$ on $\boldsymbol{M}\times\boldsymbol{C}^n$ as follows. Let $(x, v)\in\boldsymbol{M}\times\boldsymbol{C}^n$. Consider all pairs $[f, \alpha]$ consisting of a holomorphic map $f:\boldsymbol{D}\rightarrow\boldsymbol{M}$ together with a complex number α such that $f(0)=x$ and $\alpha f'(0)=v$. Then

$$k_{\boldsymbol{M}}(x, v) = \inf|\alpha|,$$

where the infimum is taken over all such maps f with numbers α. $k_{\boldsymbol{M}}$ is a biholomorphic invariant, in the following sense: if $F:\boldsymbol{M}\rightarrow\boldsymbol{N}$ is biholomorphic, then

$$k_{\boldsymbol{M}}(x, v) = k_{\boldsymbol{N}}(F(x), dF_x(v)),$$

where $dF_x:\boldsymbol{C}^n\rightarrow\boldsymbol{C}^n$ is the linear transformation defined by the matrix of partial derivatives of F at x. Note that $k_{\boldsymbol{M}}(x, cv)=|c|k_{\boldsymbol{M}}(x, v)$ for all $(x, v)\in\boldsymbol{M}\times\boldsymbol{C}^n$ and all complex numbers c.

First, we compute $k_{\boldsymbol{B}}(\boldsymbol{0}, v)$. Let v and w be two vectors in $\boldsymbol{C}^2$ such that $|v|=|w|$, and let F be a complex-linear rotation of $\boldsymbol{C}^2$ with $F(v)=w$. Clearly F maps the unit ball $\boldsymbol{B}$ biholomorphically onto itself, and $dF_0(v)=w$. Thus $k_{\boldsymbol{B}}(\boldsymbol{0}, v)=k_{\boldsymbol{B}}(\boldsymbol{0}, w)$.

This shows that $k_{\boldsymbol{B}}(\boldsymbol{0}, v)$ must depend only on $|v|$; we see that it must be $c|v|$ for some positive constant c. In fact, $c=1$, although we do not need this information. We do need to notice, however, that $k_{\boldsymbol{B}}(\boldsymbol{0}, \cdot)$ is a differentiable function on $\boldsymbol{C}^2-\{\boldsymbol{0}\}$.

We now compute $k_{\boldsymbol{D}\times\boldsymbol{D}}(\boldsymbol{0}, v)$. At first, let $v=(1, m)$ with $|m|\leqq 1$. If $f=(f_1, f_2):\boldsymbol{D}\rightarrow\boldsymbol{D}\times\boldsymbol{D}$ with $f(0)=\boldsymbol{0}$ and $\alpha f'(0)=\alpha(f_1'(0), f_2'(0))=(1, m)$, then by the Schwarz Lemma, $|f_1'(0)|\leqq 1$ and $|\alpha|\geqq 1$. Thus $k_{\boldsymbol{D}\times\boldsymbol{D}}(\boldsymbol{0}, v)\geqq 1$. On the other

hand, the value 1 for α is actually attained by the map $f(z)=(z, mz)$. Hence $k_{D\times D}(0, v)=1$, where $v=(1, m)$.

In general, if $v=(a, b)$ with $|b|\leqq|a|$, then

$$k_{D\times D}(0, v) = |a|\, k_{D\times D}(0, (1, b/a)) = |a|.$$

Finally, $k_{D\times D}(0, (a, b))=\text{Max}\{|a|, |b|\}$. But this is *not* a differentiable function on $C^2-\{0\}$. Now since the disk D is homogeneous, so is the bidisk $D\times D$ and $k_{D\times D}(x, \cdot)$ is not differentiable on $C^2-\{0\}$ for any $x\in D\times D$. Thus no biholomorphic map from B to $D\times D$ can exist.

The author is supported in part by NSF grant GP-12403.

References

1. H. Behnke, and P. Tullen, Theorie der Funktionen mehrerer komplexer Veränderlichen, Springer-Verlag, Berlin, 1934.

2. S. Kobayashi, Invariant distances on complex manifolds and holomorphic mappings, J. Math. Soc. Japan, 19 (1967) 460–480.

SEPARATELY CONTINUOUS FUNCTIONS ARE BAIRE FUNCTIONS

F. W. CARROLL, The Ohio State University

This theorem is well known, and has been much generalized [1], but I find the following proof amusing.

THEOREM. *Let f be a real function on the unit square which is continuous in each variable separately. Then f is the pointwise limit of a sequence of jointly continuous functions.*

Proof. We may assume that f is bounded (otherwise, consider arc tan f) and that $f(0, y)=f(1, y)$ for each y (otherwise, consider $f(x, y)-x(f(1, y)-f(0, y))$. For $K=1, 2, \cdots$, let

$$S_K(x, y) = \sum_{k=-K}^{K}\left(1-\frac{|k|}{K}\right)a_k(y)e^{2\pi ikx}, \tag{1}$$

where

$$a_k(y) = \int_0^1 f(x, y)e^{-2\pi ikx}dx \qquad (k = 0, \pm 1, \cdots). \tag{2}$$

For fixed y, $S_K(x, y)$ converges (uniformly in x) to $f(x, y)$ as K increases, by Fejer's theorem. It remains to prove that each a_k is continuous. If $\{y_\nu\}$ is a sequence converging to y_0, then by (2), $a_k(y_\nu)-a_k(y_0)$ is a sequence of integrals. From the continuity of f in y, the sequence of integrands converges pointwise to zero. By the bounded convergence theorem, the sequence of integrals converges to zero, so that a_k is continuous at each y_0.

References

1. K. Kuratowski, Topology I, Academic Press, New York, 1966, p. 378.

ON PLANAR GRAPHS

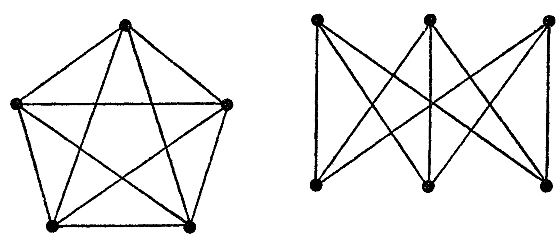
T. D. PARSONS, Pennsylvania State University

1. Introduction. Our purpose is to use the well-known theorem of Kuratowski [4] characterizing planar graphs to obtain an easy step-by-step proof of Whitney's theorem [6] that only planar graphs have duals. Since Kuratowski's theorem is usually proved in graph theory texts (see [1], [2], [3], [5]) but Whitney's theorem usually stated without proof, our proof will be of interest to students of graph theory.

Let us use the definition that two graphs are **duals** if there is a bijection between their sets of edges such that the polygons of one graph correspond to the cut-sets of the other, and vice versa. Here polygon means a simple closed path with no proper closed subpath, and **cut-set** means a minimal set of edges whose removal increases the number of connected components. Liu [5, p. 222] has introduced this definition, which is easily shown to be equivalent to Whitney's original definition of dual graphs in [6]. We may now state:

WHITNEY'S THEOREM: *A graph is planar if and only if it has a dual.*

KURATOWSKI'S THEOREM: *A graph is nonplanar if and only if it contains a subgraph which is isomorphic up to vertices of degree 2 with one of the two Kuratowski graphs* (in Fig. 1).



K_5 $K_{3,3}$

FIG. 1

Two graphs are **isomorphic up to vertices of degree 2** if they are isomorphic after possible additions or suppressions of some vertices of degree 2. For example the graphs of Fig. 2 are so related to K_5 and $K_{3,3}$. (We could equally well say that these graphs are "homeomorphic" to K_5 and $K_{3,3}$.)

2. Outline of the proof. It is easy to show that planar graphs have duals: given any representation of the graph G in the plane, we construct a dual graph G^* in the usual way, by putting a vertex of G^* in the interior of each region bounded by edges of G in the plane, then connecting vertices of G^* in adjacent regions by edges cutting across the edges of G in a one-one correspondence. To show that nonplanar graphs do not have duals, we verify six assertions:

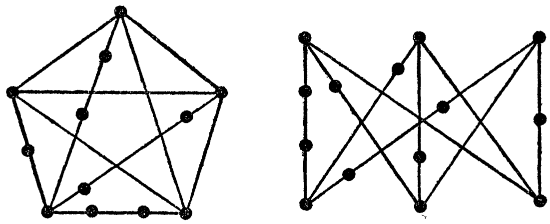


FIG. 2

(1) The graphs K_5 and $K_{3,3}$ have no duals.

(2) If a graph G has a dual G^*, and if e is an edge of G, then the graph H, obtained from G by deleting the edge e, has a dual H^*.

(3) If G has a dual, then so does each subgraph of G.

(4) If G has a dual G^* and H is obtained from G by adding or suppressing a vertex of degree 2, then H has a dual H^*.

(5) If G has a dual, then so does each graph which is isomorphic to G up to vertices of degree 2.

(6) (by Kuratowski's Theorem): If G is nonplanar, then G has no duals.

3. Verification of the assertions. (1) Suppose G is a dual of $K_{3,3}$. Since $K_{3,3}$ has 9 edges, no cut-sets of 2 edges, and polygons of lengths 4 and 6 only, G must have 9 edges, no polygons of length 2, and vertices of degree 4 or more. Thus G has at least 5 vertices, and so at least $\frac{1}{2}(5)(4)=10$ edges, a contradiction.

Suppose G is a dual of K_5. Since K_5 has 10 edges, no polygons of length 2, and cut-sets with 4 or 6 edges only, G must have 10 edges, no vertices of degree 2, and polygons of lengths 4 and 6 only. (Thus G is bipartite, for all its closed paths are of even length.) A bipartite graph on 6 or fewer vertices would have at most 9 edges; thus G must have at least 7 vertices, hence at least $\frac{1}{2}(7)(3)>10$ edges, a contradiction.

(2) Let e^* be the edge of G^* corresponding to the edge e of G. Let H^* be obtained from G^* by contracting e^* to a point. It is easy to verify that H^* is a dual of H.

(3) Since each subgraph of G can be obtained by successive deletions of single edges, this follows from (2).

(4) If H arises from G by introducing a vertex of degree 2 on some edge e of G, and if e^* is the corresponding edge of G^*, let H^* be obtained from G^* by adding a new edge with the same vertex-ends as e^* in G^*. (The new edge is "parallel" to the edge e^* in G^*.)

If H instead arises from G by suppression of a vertex v of degree 2 in G, then the two edges e_1 and e_2 incident to v in G have "parallel" correspondents e_1^* and e_2^* in G^*, and we obtain H^* from G^* by deleting e_2^*.

In either case, we may easily verify that H and H^* are duals.

(5) This clearly follows from (4).

(6) If G is nonplanar, then by Kuratowski's theorem, G has a subgraph H

isomorphic up to vertices of degree 2 with K_5 or $K_{3,3}$. If G had a dual, so would H by (3), thus so would K_5 or $K_{3,3}$ by (5). This contradicts (1). Therefore G has no duals.

The author is indebted to one of his students, Mr. S. Axler of Princeton University, for suggesting the insertion of (2) before (3), which could be proved directly. This in turn suggested the insertion of (4) before (5). The intermediate steps seem to make the proof clearer to beginning students of graph theory.

References

1. C. Berge, The Theory of Graphs, Wiley, New York, 1966, Chapter 21.
2. R. G. Busacker and T. L. Saaty, Finite Graphs and Networks, McGraw-Hill, New York, 1965, pp. 66–73.
3. F. Harary, Graph Theory, Addison-Wesley, Reading, Mass., 1969, pp. 108–116.
4. C. Kuratowski, Sur le Problème des Courbes Gauches en Topologie, Fund. Math., 15 (1930) 271–283.
5. C. L. Liu, Introduction to Combinatorial Mathematics, McGraw-Hill, New York, 1968, pp. 211–225.
6. H. Whitney, Nonseparable and planar graphs, Trans. Amer. Math. Soc., 34 (1932) 339–362.

RESEARCH PROBLEMS

EDITED BY RICHARD GUY

In this Department the Monthly presents easily stated research problems dealing with notions ordinarily encountered in undergraduate mathematics. Each problem should be accompanied by relevant references (if any are known to the author) and by a brief description of known partial results. Manuscripts should be sent to Richard Guy, Department of Mathematics, Statistics, and Computing Science, The University of Calgary, Calgary 44, Alberta, Canada.

IF A IS A SYMMETRIC *-ALGEBRA IS A_e SYMMETRIC?

R. S. DORAN, Texas Christian University

A **-algebra* is a complex associative linear algebra A with a mapping $x \rightarrow x^*$ of A into itself such that for $x,y \in A$ and complex λ: (a) $(x+y)^*=x^*+y^*$; (b) $(xy)^*=y^*x^*$; (c) $(\lambda x)^*=\bar{\lambda}x^*$($\bar{\lambda}$ is the complex conjugate of λ); and (d) $x^{**}=x$. The map $x\rightarrow x^*$ is called an *involution*; because of (d) it is clearly bijective.

Typical examples of *-algebras are the complex numbers C with involution $\lambda^*=\bar{\lambda}$ (complex conjugation); the algebra $\mathcal{C}(X)$ of bounded continuous complex-valued functions on a topological space X with involution $f^*(t)=\overline{f(t)}$; the algebra $\mathcal{B}(H)$ of bounded linear operators on a Hilbert space H with involution $T\rightarrow T^*$ (the adjoint of T); the group algebra $L^1(G)$ of a locally compact (non-discrete) abelian group G with involution $f^*(t)=\overline{f(-t)}$; and the algebra $\mathcal{C}(D)$ of continuous complex-valued functions on the closed unit disc analytic on the interior of D with involution $f^*(\lambda)=\overline{f(\bar{\lambda})}$.

A*-algebra may or may not possess an identity. In the above examples only the group algebra $L^1(G)$ fails to have an identity. If A is a *-algebra without identity, a new *-algebra A_e may be constructed from A in a standard way which contains A as a maximal two-sided ideal, does have an identity, and which reflects most of the properties that A has. Indeed, one sets $A_e = \{(x, \lambda) : x \in A, \lambda \in C\}$ and introduces the following operations on A_e:

$$\alpha(x, \lambda) = (\alpha x, \alpha\lambda), \quad (x, \lambda) + (y, \mu) = (x + y, \lambda + \mu),$$

$$(x, \lambda)(y, \mu) = (xy + \lambda y + \mu x, \lambda\mu), \quad \text{and} \quad (x, \lambda)^* = (x^*, \bar{\lambda}).$$

An alternative method for handling *-algebras A without identity is to consider the circle operation $x \circ y = x + y - xy$ for $x, y \in A$. An element x in A is said to be *quasi-regular* if there exists y in A such that $x \circ y = 0 = y \circ x$; the element y is unique, and is called the *quasi-inverse* of x. The set of quasi-regular elements in A is a group relative to the circle operation (with identity 0). If A has an identity e, the motivation for introducing the circle operation can quickly be seen by examining the simple relation $e - (x \circ y) = (e - x)(e - y)$.

A *-algebra A is said to be *symmetric* if for each x in A the element $-x^*x$ has a quasi-inverse. If A has an identity e, this means (utilizing the relation above) that $e + x^*x$ has an inverse for each x in A. Many important *-algebras are symmetric; for instance, all of the examples given above except the last one are symmetric.

PROBLEM: *If A is a symmetric *-algebra without identity, is the *-algebra A_e obtained from A by adjoining an identity symmetric?*

The reader will note that the problem is purely algebraic, and that its statement depends only on the few simple definitions given above. The answer is known to be affirmative if A is a Banach *-algebra [2, page 233]. In this case powerful representation techniques from the theory of Banach algebras may be applied. Little else appears to be known even in the presence of commutativity. Because of this, a solution to the above problem would be of considerable interest. A good summary of *-algebras and their properties can be found in [1] and [2].

References

1. M. A. Naimark, Normed Rings, 2nd ed., Noordhoff, Groningen, the Netherlands, 1964.
2. C. E. Rickart, General Theory of Banach Algebras, Van Nostrand, Princeton, N. J., 1960.

WHICH POLYNOMIALS OVER AN ALGEBRAIC NUMBER FIELD MAP THE ALGEBRAIC INTEGERS INTO THEMSELVES?

D. A. LIND, Stanford University

This problem originally arose while trying to decide the following question: If two polynomials f and g with coefficients in the integers Z are given such that

THE NODAL NUMBER OF A DOMAIN

KURT KREITH, University of California, Davis

A string of length l which is tied down at both end points is well known to have a sequence of fundamental modes of oscillation. If the string is of unit density, then these fundamental modes can be found by solving the eigenvalue problem

$$
\begin{aligned}
&u'' + \lambda u = 0; \qquad 0 < x < \\
&u(0) = u(l) = 0.
\end{aligned} \tag{1}
$$

This problem has nontrivial solutions for a sequence of values of λ called **eigenvalues**, and a nontrivial solution $\phi_k(x)$ of (1) corresponding to an eigenvalue λ_k is called an **eigenfunction**.

For the problem (1) it is readily verified that $0<\lambda_1<\lambda_2 \cdots$, that $\lim_{k\to\infty} \lambda_k = \infty$, and that the eigenfunction ϕ_k corresponding to λ_k has $k-1$ zeros in the open interval $(0, l)$. Thus the zeros of ϕ_k divide $(0, l)$ into k segments called **nodal domains** of ϕ_k.

If we generalize these considerations to n-dimensional homogeneous membranes in place of strings, we are led to the eigenvalue problem

$$
\begin{aligned}
\Delta u + \lambda u &= 0 \quad \text{in } D \\
u &= 0 \quad \text{on } B,
\end{aligned} \tag{2}
$$

where D is a smooth bounded domain in E^n with boundary B. It can be shown [1], that the eigenvalue problem (2) has nontrivial solutions for a discrete sequence of eigenvalues $0<\lambda_1<\lambda_2\leqq\lambda_3\leqq \cdots$ and that $\lim_{k\to\infty}\lambda_k=\infty$. However, the question of the number of nodal domains of the eigenfunction ϕ_k corresponding to λ_k is somewhat more complex and interesting.

A subdomain $D'\subseteq D$ is called a **nodal domain** for an eigenfunction ϕ_k if $\phi_k(x)\neq 0$ for x in D' but $\phi_k(x)=0$ for all x in the boundary of D'. Since $\phi_1(x)$ can be shown to be nonzero in D, D itself is the one and only nodal domain for ϕ_1.

We shall be interested in the number of nodal domains of the eigenfunction $\phi_k(x)$ for $k>1$ and denote this positive integer by $N(k)$. As observed above, in the case of a string $N(k)=k$ for all k. Experimentation with (2) in simple geometries shows that we cannot expect $N(k)\equiv k$ for $n>1$. However a celebrated theorem due to Courant [1] asserts that $N(k)\leqq k$ for $n\geqq 1$!

These considerations lead us to consider the **nodal number** of a domain which is defined as follows: The nodal number of a domain D is given by

$$\eta(D) = \sup\{k \mid N(j) = j \quad \text{for } j = 1, 2, \cdots, k\}.$$

From classical results observed above and developed in [1], the following properties of $\eta(D)$ follow easily:

1. In E^1 (2) reduces to (1) and $N(k)=k$ for all k. Therefore in E^1, $\eta(D)=+\infty$ for any domain (a, b).

2. In E^n, $N(1)=1$ and $N(2)=2$. Therefore $\eta(D)\geqq 2$ for any domain D. However if D is a square in E^2, then $N(3)=2$ and $\eta(D)=2$.

A less obvious property of $\eta(D)$ follows from a theorem due to A. Pleijel [2]. This theorem asserts that if $n\geqq 2$, then

$$\{k \mid N(k) = k\}$$

is a finite set. Therefore,

3. If $n\geqq 2$, then $\eta(D)<\infty$.

The question arises as to whether there are domains in $E^n(n\geqq 2)$ with large nodal numbers and what such domains look like. In this connection it is of interest to consider the rectangle

$$D = \{(x,y) \mid 0 < x < a\pi,\ 0 < y < b\pi\}$$

in E^2. The eigenfunctions of (2) are then given by

$$\phi(x) = \sin\frac{ix}{a}\sin\frac{jy}{b}; \qquad i,j = 1, 2, \cdots$$

and the corresponding eigenvalues by $\lambda=(i^2/a^2)+(j^2/b^2)$. Suppose $a>b$ so that

$$\lambda_1 = \frac{1}{a^2}+\frac{1}{b^2}; \qquad \lambda_2 = \frac{4}{a^2}+\frac{1}{b^2}\cdot$$

The question of whether $\eta(D)\geqq 3$ or $\eta(D)<3$ is decided by whether or not

$$\frac{9}{a^2}+\frac{1}{b^2}\leqq\frac{4}{a^2}+\frac{4}{b^2}.$$

More generally, $\eta(D)=k$ if

$$\frac{k^2}{a^2}+\frac{1}{b^2}\leqq\frac{(k-1)^2}{a^2}+\frac{4}{b^2} \quad\text{and}\quad \frac{(k+1)^2}{a^2}+\frac{1}{b^2}>\frac{k^2}{a^2}+\frac{4}{b^2}\cdot$$

By choosing a much larger than b, it is possible to make $\eta(D)$ arbitrarily large.

These considerations suggest that it is long narrow domains which have large nodal numbers and they raise the following question: Given a domain D, is there some measure of being "long and narrow" which yields bounds for $\eta(D)$?

The answer is "yes" for rectangles in E^2 by the above computations. A similar computation is possible for rectangles in E^n for $n>2$. For a convex set D a measure of being "long and narrow" would be the ratio between a diameter and the largest distance between supporting planes parallel to that diameter. It would be of interest to obtain bounds for $\eta(D)$ in terms of these or other parameters.

References

1. R. Courant and D. Hilbert, Methods of Mathematical Physics, Vol. 1, Interscience, New York, 1953.

2. A. Pleijel, Remarks on Courant's nodal line theorem, Comm. Pure Appl. Math., 9 (1956), 543–550.

CLASSROOM NOTES

Manuscripts for this Department should be sent to David Drasin, Division of Mathematical Sciences, Purdue University, Lafayette, IN 47907.

MY FAVORITE PROOF OF MEHLER'S INTEGRAL

PETER HENRICI, Eidgenössische Technische Hochschule, Zürich

In a recent issue, Askey [1] suggested a proof of Mehler's integral representation of the Legendre polynomial. As it stands, Askey's proof is incomplete, because it assumes the convergence of the generating series on the boundary of the disk of convergence. Also, the simplicity claimed for the proof is, perhaps, deceptive, because the proof appeals to a principle that a series whose sum is integrable may be integrated term by term. Although this is true for Fourier series, it does not hold for series in general.

Askey's idea could be turned into a correct elementary proof by substituting $r=(1-\epsilon)e^{i\phi}$ in the generating series, where $\epsilon>0$, and letting $\epsilon\to 0$ in the final result, but the labor required for justifying the interchange of a limit and an improper integration would tend to deflate further the author's claim of simplicity. In my own teaching practice I prefer a proof which links Legendre polynomials to Laplace's equation and uses some basic complex variable theory. My proof is based on a simple Lemma which simultaneously serves as a key to other integral representations of special functions of mathematical physics.

Let (x, y, z), (r, ϕ, z), (R, ϕ, θ) be cartesian, cylindrical, and spherical coordinates in R^3, respectively, so that

$$(1)\qquad \begin{cases} x = r\cos\phi = R\sin\theta\cos\phi,\\ y = r\sin\phi = R\sin\theta\sin\phi,\\ \qquad\quad\; z = R\cos\theta, \end{cases}$$

$(0\leqq\phi<2\pi,\ 0\leqq\theta\leqq\pi)$. In all that follows, these coordinates are real. We consider complex-valued harmonic functions, i.e., solutions of Laplace's equation

$$(2)\qquad \frac{\partial^2 u}{\partial x^2}+\frac{\partial^2 u}{\partial y^2}+\frac{\partial^2 u}{\partial z^2}=0,$$

that are symmetric with respect to the z-axis. Such solutions depend only on r and z, or on R and θ. If the domain of definition of such a solution u intersects the z-axis, and if the restriction of u to the z-axis is a real-analytic function of z, we shall call u a *regular axially symmetric harmonic function.*

LEMMA [2, 3, 5]. *Let S be a convex domain in the complex plane which is symmetric with respect to the real axis. Let the function $f: t\to f(t)$ $(t\in S)$ be holomorphic. Then there exists, in the domain of all (r, z) such that $z+ir\in S$, a unique regular axially symmetric harmonic function $u:(r, z)\to u(r, z)$ such that $u(0, z)=f(z)$ for z real, $z\in S$. This function is represented in terms of f as follows:*

$$u(r, z) = \frac{1}{\pi}\int_0^\pi f(z + ir\cos\phi)d\phi. \tag{3}$$

Proof. The uniqueness statement is proved by a straightforward power series argument, using Laplace's equation in cylindrical coordinates. Differentiating under the integral sign and integrating by parts one then can verify that the function (3) has all the required properties.—The following more elegant argument (communicated to me by Alfred Huber who learned it from Marcel Riesz) shows how (3) might have been discovered: Clearly, $(x, y, z)\to f(z+ix)$ is harmonic. But then, by the invariance of Laplace's equation under rotations, so is $(x, y, z)\to f(z+i(x\cos\alpha+y\sin\alpha))$ for every fixed angle α. All these functions for $r=0$ equal $f(z)$ and hence are real-analytic on the z-axis. Their average (3) clearly is harmonic, and axially symmetric.

APPLICATION 1. Let n be a nonnegative integer, and denote by P_n the Legendre polynomial. Separating Laplace's equation in spherical coordinates we find that $(r, z)\to R^nP_n(\cos\theta)$ is a regular axially symmetric harmonic function which for $r=0$ equals z^n. Thus by the lemma,

$$R^nP_n(\cos\theta) = \frac{1}{\pi}\int_0^\pi (z + ir\cos\phi)^n d\phi$$

or, by virtue of (1),

$$P_n(\cos\theta) = \frac{1}{\pi}\int_0^\pi (\cos\theta + i\sin\theta\cos\phi)^n d\phi.$$

This is *Laplace's integral* for P_n.

APPLICATION 2. Separating Laplace's equation in cylindrical coordinates we obtain (among others) the regular axially symmetric harmonic function $(r, z)\to J_0(r)e^z$, where J_0 is the Bessel function of order zero. Here $u(0, z)=e^z$. Thus the lemma (with $z=0$) yields the *Parseval-Bessel-Poisson integral*

$$J_0(r) = \frac{1}{\pi}\int_0^\pi e^{ir\cos\phi}\, d\phi.$$

APPLICATION 3. The integral (3) may be viewed as a complex line integral. Letting $t=z+ir\cos\phi$ we obtain

$$u(r, z) = \frac{1}{i\pi}\int_{z-ir}^{z+ir} [r^2 + (t - z)^2]^{-1/2} f(t)dt. \tag{4}$$

The square root is positive for t real, which requires a cut in the t-plane, say, from $z+ir$ to $z+i\infty$ and a symmetric cut in the lower half-plane. Otherwise, the path of integration is arbitrary in S. We consider once more the solution $u(r, z)=R^nP_n(\cos\theta)$ and choose as path of integration the circular arc $t=Re^{i\phi}$, $-\theta\leq\phi\leq\theta$. Then $dt=iRe^{i\phi}d\phi$,

$$r^2 + (t - z)^2 = 2Re^{i\phi}[\cos\phi - \cos\theta],$$

and we immediately obtain *Mehler's integral,*

$$P_n(\cos\theta) = \frac{1}{\pi}\int_{-\theta}^{\theta} e^{i(n+1/2)\phi}[2\cos\phi - 2\cos\theta]^{-1/2}d\phi.$$

We leave it to the reader to use the method to (re-)discover other integral representations for known solutions of Laplace's equation. The scope of the method is not restricted to axially symmetric solutions, since integral operators similar to (3) are known to exist for other types of solutions, and also for other partial differential equations [3, 4].

References

1. R. Askey, Mehler's integral for $P_n(\cos\theta)$, this MONTHLY, 76 (1969) 1046–1049.

2. A. Erdélyi, Singularities of generalized axially symmetric potentials, Comm. Pure Appl. Math., 9 (1956) 403–414.

3. P. Henrici, Zur Funktionentheorie der Wellengleichung, Comm. Math. Helv., 27 (1953) 235–293.

4. ———, Complete systems of solutions for a class of singular elliptic partial differential equations. Boundary problems in differential equations, R. E. Langer (ed.), The University of Wisconsin Press, 1960, pp. 19–34.

5. G. Szegö, On the singularities of zonal harmonic expansions. J. Rational Mech. Anal.. 3 (1954) 561–564.

ON THE CONNECTED SUM OF PROJECTIVE PLANES, TORI, AND KLEIN BOTTLES

J. C. ALEXANDER, Johns Hopkins University and University of Maryland

The classification of compact closed surfaces is a standard chore in introductory algebraic topology courses. In Massey [1], the classification is stated in Theorems 5.1 and 7.2 of chapter one. A step in this classification is the following:

PROPOSITION. 1. *The connected sum of two projective planes is homeomorphic to a Klein bottle.*

2. *The connected sum of a Klein bottle and a projective plane is homeomorphic to the connected sum of a torus and a projective plane.*

We offer here a proof of this proposition alternative to that of [1]; it is shorter and more mechanical.

Let P = projective plane, T = torus, K = Klein bottle, $\#$ denote connected sum.

The Klein bottle is well known to be representable as a square with identifications as in Figure 1 below. Thus the two edges a are to be identified with (glued to) each other with the orientations indicated by the arrows, and likewise the edges b. The connected sum $P \# P$ is also represented. (See [1] chap. 1, fig. 1.9.)

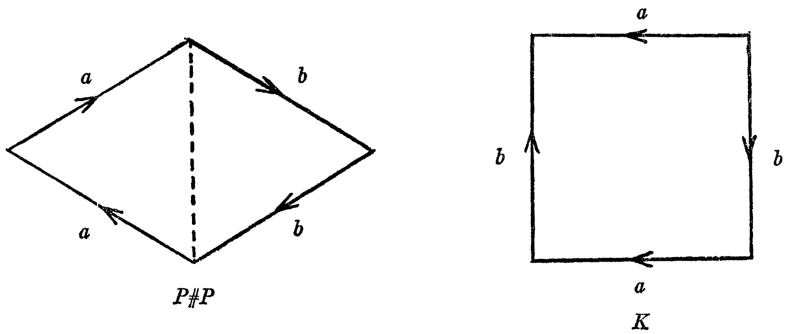


FIG. 1

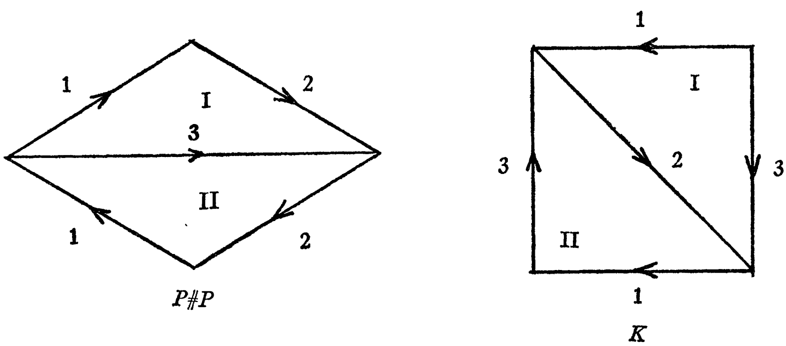


FIG. 2

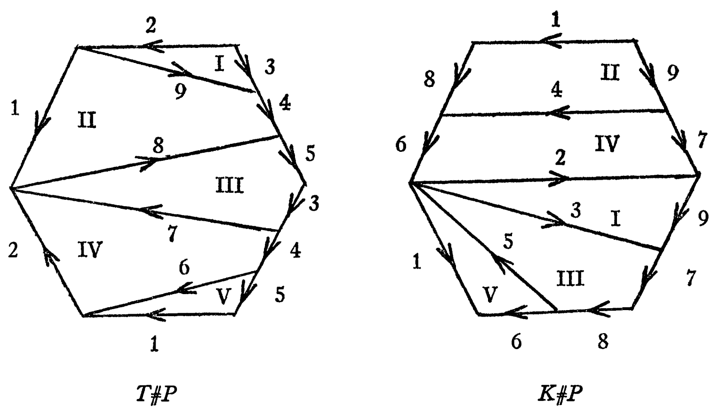


FIG. 3

Figure 2 shows these to be homeomorphic. It is interpreted as follows. We take $P \# P$ and cut it into two triangles I, II by cutting along the indicated lines. These triangles are then rearranged and glued back together to get K. The

numbers indicate the edges. It is simple to check that everything fits together properly.

In the same way, Figure 3 shows that $T \# P = K \# P$. The P in each case is associated with the right hand two edges of the hexagon.

Reference

1. W. S. Massey, Algebraic Topology: An Introduction, Harcourt, Brace and World, New York, 1967.

PATHOLOGY OF UPPER STONE-ČECH COMPACTIFICATIONS

S. Salbany and G. C. L. Brümmer,
University of Cape Town (South Africa)

This note discusses a type of compactification of which a special case was introduced in this Monthly under the designation "upper Stone-Čech compactification" by Nielsen and Sloyer [5].

Let $R^\bullet$ denote the space of the reals with the upper topology, that is, having the open intervals of the form $(-\infty, x)$ as base for open sets. Any subset A of R, when equipped with the relative topology from $R^\bullet$, will be denoted by $A^\bullet$.

Let $A^\bullet$ have at least two points. Each topological space X is initial for the set $C(X, A^\bullet)$ of all continuous mappings $X \to A^\bullet$, that is, X has the coarsest topology making the mappings continuous. These mappings distinguish points if and only if X is a T_0-space. Consider the mapping $e_{A,X}$ (for brevity, e) of X into the topological product

$$A^{\bullet C(X,A^\bullet)} \tag{1}$$

given by $\pi_\alpha \circ e = \alpha$ for all $\alpha \in C(X, A^\bullet)$. It is now immediate from considerations of initiality [6, p. 781] that e is a topological embedding if and only if X is a T_0-space. We denote the closure of the image $e[X]$ in the product (1) by $\beta_A^\bullet X$. The pair $(e, \beta_A^\bullet X)$ is a T_0-compactification of X when $A^\bullet$ is compact and X is T_0.

Nielsen and Sloyer [5] considered the case $A = [0, 1] = I$, say, and asserted that if $f: X \to Y$ is continuous, X being T_0 and Y compact Hausdorff, then there exists a unique continuous $\bar{f}: \beta_I^\bullet X \to Y$ such that $\bar{f} \circ e = f$. No proof was given but it was claimed that "the usual methods" as in Kelley [4] yielded a proof. However, the form of proof in Kelley would only furnish an $\bar{f}$: $\beta_I^\bullet X \to \beta_I^\bullet Y$ with $\bar{f} \circ e_{I,X} = e_{I,Y} \circ f$. (Indeed, the proof in [4 p. 153] has a gap, pointed out to one of us by G. H. Toomer in 1967.) To settle the matter, we prove the following result:

Theorem. *Let X be a topological space, Y a T_1-space and $A^\bullet$ compact. Then every continuous $g: \beta_A^\bullet X \to Y$ is constant.*

Proof. By its compactness, $A^\bullet$ has a rightmost point r. In the product (1), consider the point b, each of whose coordinates is r. Any neighborhood of b in

the product contains an intersection of finitely many sets $\pi_\alpha^{-1}[V_\alpha]$, with V_α a neighborhood of r in $A^\bullet$, whence $V_\alpha = A^\bullet$. Thus the only neighborhood of b is the whole product, and so b belongs to each nonvoid closed subset of the product. Now for continuous $g:\beta_A^\bullet X \to Y$ and any $x \in \beta_A^\bullet X$, $g^{-1}[g(x)]$ is nonvoid and closed in the product. Hence $b \in g^{-1}[g(x)]$. Thus $g(b) = g(x)$, and g is constant.

COROLLARY. *If Y is a T_1-space (a fortiori, if Y is compact Hausdorff), then for nonconstant continuous $f:X \to Y$ there is no continuous $\bar{f}:\beta_I^\bullet X \to Y$ with $\bar{f} \circ e = f$.*

As a further corollary we have the fact, implicit in [1 p. 508] for the case $A = \{0, 1\}$, that $(e, \beta_A^\bullet X)$ is not a compact reflection in the category of T_0-spaces.

REMARK. The topological properties of the compactification $(e, \beta_A^\bullet X)$ depend on $A^\bullet$; e.g. if $\{0, 1\}$ and $[0, 1]$ respectively are substituted for A, then topologically inequivalent compactifications are obtained [3]. Further, there are analogies between the quasi-uniform properties of $(e, \beta_A^\bullet X)$ and the uniform properties of the Stone-Čech compactification. These analogies are entailed by ([2] Theorem 1.2 and Proposition 1.5) and depend on the quasi-uniform structure given to $A^\bullet$ [3].

The first author was supported by a Messina Company fellowship. The second author's work was supported in part by a University of Cape Town Staff Research Grant.

References

1. S. Baron, Reflectors as compositions of epireflectors, Trans. Amer. Math. Soc., 136 (1969) 499–508.

2. G. C. L. Brümmer, Initial quasi-uniformities, Nederl. Akad. Wetensch. Proc. Ser. A, 72 = Indag. Math. 31 (1969) 403–409.

3. ———, On the quasi-uniform totally bounded reflection, (in preparation).

4. J. L. Kelley, General Topology, Van Nostrand, New York, 1955.

5. R. Nielsen and C. Sloyer, On embedding in quasi-cubes, this MONTHLY, 75 (1968) 514–515.

6. J. C. Taylor, Weak families of maps, Canad. Math. Bull., 8 (1965) 771–781

MATHEMATICAL EDUCATION

EDITED BY J. G. HARVEY AND M. W. POWNALL

Material for this Department should be sent to either of the editors: J. G. Harvey, Department of Mathematics, University of Wisconsin, Madison, WI 53706; M. W. Pownall, Department of Mathematics, Colgate University, Hamilton, NY 13346.

ON THE NOTION OF "FUNCTION"

G. J. MINTY, Indiana University

I recently became suspicious that my students in Differential Equations did not fully appreciate the existence- and uniqueness-theorems for the usual initial-value problem for $y' = f(x, y)$ because they didn't understand what f is. I asked

them to define "function," and they (at least, some of them) gave me the standard "set of ordered pairs" definition. Pushing the point further, I asked on a quiz: Which of the following can be interpreted as differential equations of the form $f(y', y, x) \equiv 0$:

$$(1)\ \int_0^{y'} e^{-(xt)^2}dt = 0; \quad (2)\ \int_0^1 \{[y'(x)]^2 + y(x)\}dx = 0; \quad (3)\ y'(y(x)) = 0?$$

The answers I received were completely random. All this makes me feel like a grade-school teacher asking his pupils "What is an integer?" and getting back Peano's axioms, and then discovering his pupils can't count to ten.

I conclude that the "set of ordered pairs" definition is doing our students even less good than the old "a number which jumps up and down while another number is jumping up and down" that I learned as an undergraduate. Certainly the ultimate test of whether the students have properly absorbed the concept is whether they can recognize a function when they see one.

It is my feeling that calculus texts and teachers ought to be doing something which they are now doing very inadequately: giving the student *many different* ways to visualize the concept (e.g.: the graph; the collection of strings tying points of the domain to points of the range; the idea that if $f(x)$ is x^2+e^x then f is $(\cdot)^2+e^{(\cdot)}$; the "slot-machine" into which, when one inserts a number and turns the crank, one gets out another determined entirely by the first; etc.), reserving the "ordered pairs" definition until after the student has assimilated these mental pictures and made his peace with them. The two virtues of the "ordered pairs" definition are its precision—which the student cannot appreciate until he sees how it ties together all these foggy mental pictures—and the demonstration that "function" can be defined within the framework of set-theory—which could well be postponed to a later course in axiomatic set-theory, and *certainly* could be postponed until the student can see for himself the set of ordered pairs associated with $f(y', y, x) = \int_0^{y'} e^{-(xt)^2}\, dt$.

EVALUATION OF ELEMENTARY TEXTBOOKS

C. B. ALLENDOERFER, University of Washington

1. Introduction. How to choose elementary textbooks is one of the most frustrating problems of our profession, for there is no satisfactory yardstick of quality. I have sat on text selection committees off and on for over 35 years and have observed all kinds of irrational behavior. Texts are chosen or rejected on the personality of the author; one text was rejected because the margins were too wide; another because of the color of the cover. Texts are judged to be too small or too large, too wordy or too brief, too easy or too hard. The real criterion of how they will affect our students is seldom used because we have no information on this.

Descriptions by the publishers represent the creative writing of their adver-

tising departments with claims such as: "This unparalleled new text," "This text is outstanding," "Broad coverage of ideas and techniques," " . . . appeals directly to students in all disciplines," "A new and distinctive approach," and so on *ad nauseam*. (The above quotes are from a recent issue of the MONTHLY.) So we adopt almost at random and change from bad to worse every few years.

This situation is surprising (even in the days of Madison Avenue) when we consider how other commodities are marketed. The usual advertisements attempt to tell us what their product will do for the buyer: "These tires will last 40,000 miles," "This toothpaste will reduce cavities by 20%," "This razor will eliminate 5 o'clock shadow"; "This car has 300 horsepower." These claims may be exaggerated, but they are at least quantitative and are subject to verification by consumers' organizations.

Not so with textbooks. No quantitative claims are ever made as to how well they teach, or what are their effects on students. They are sometimes "classroom tested," but no data are ever given so that a potential user can judge whether these tests showed a success or a failure. So we choose books as blind men in the dark.

The same problem holds for conscientious authors. How can they decide what approach works best for students, how can they test their exposition, and so how can they judge their own work and improve it?

Being fed up with lack of answers to these questions, I have conducted a quantitative experiment in evaluation and revision which is described in this article. The results were so gratifying that I urge others to follow similar paths in the hope that in time we may be able to choose our textbooks in a far more rational way than is done at present.

2. The experiment. The key to the process of evaluation is the construction of "behavioral objectives" (psychological jargon). Simply this means that corresponding to each small section of the text one writes down very explicitly what the student should be able to do after reading that section and working the corresponding exercises. For example: "By completing the square to be able to solve quadratic equations which have real roots and whose coefficients are integers between -20 and 20." Associated with each stated objective is a short set of problems of the type described in the objective that are to be used to test whether the text teaches this objective effectively.

In the first instance I used this method to evaluate the preliminary edition of a text that I had written in association with others. In some places it was shocking to find that the text was so muddy that I could not discover what the objectives really were. This is a humbling experience for an author. But for the most part, the objectives could be discovered with reasonable ease. From this list of objectives and associated problems and questions, short daily tests were prepared. The students were told to read a given part of the text, and a test on this part was administered in the first ten minutes of the next class meeting. After the test, the class session was open for questions and discussion, but no lecture

on advance material was given. Thus the test scores gave an evaluation of the text independent of the instructor. When the course was over, I had remarkably specific information on weaknesses in the text, most of which could be pinpointed to single bad sentences, to insufficient worked examples, or to inadequate exercise material.

Before I gave the course again the two chapters with the worst record were completely revised in the hope that their weaknesses were removed, and smaller changes were incorporated in other chapters. Then the cycle of testing was repeated. To my great pleasure the student performance on the second go-round was significantly improved in nearly all respects.

It is impossible in a short article to give the full details of such an experiment, and so I shall not do so. The full set of objectives, problems, and scores together with a statistical analysis of the data has been published in a pamphlet titled *Evaluation and Revision of Text Materials* (273 pages) which may be obtained from me on request.

3. How future texts should be written and marketed. As a result of this experience I believe that I have some valid suggestions for future authors, publishers, and users.

(a) In writing a text the author should simultaneously prepare a list of behavioral objectives and associated test items. This will be of great assistance to him in organizing the material, and in improving the clarity of the writing.

(b) Before any manuscript is published it should be tested in a preliminary edition by the method described above. If the scores are substandard, one or more revisions and associated testings may be required. If the scores cannot be brought to a satisfactory level, the manuscript should be abandoned.

(c) When the text is published, it should be accompanied by a manual including the list of objectives and test items, the item scores obtained on the tests of the latest revision, and a description of the abilities and backgrounds of the students tested. Testing on different ability groups of students would be highly desirable.

(d) Then the publisher can advertise that this text will result in a specific level of performance on this set of objectives for specific bodies of students.

(e) From among the books so published, the user can then select that one which most nearly fits his own objectives for his course and which will result in an acceptable level of performance on the part of his students. At last he will have a good idea of what he is buying and no longer will be wandering in the jungle of advertising blurbs and uninformed impressions.

It is evident that this process will require several times the effort now expended by authors on their manuscripts and that development costs for publishers will escalate. If, however, this results in the publication of a limited number of effective texts instead of the current multitude of pot-boilers, all this effort and money will have been well spent. For the good of our students we cannot afford to do otherwise.

WHY CAN'T WE HAVE A USA MATHEMATICAL OLYMPIAD?

NURA D. TURNER, State University of New York at Albany

Requests for the questions of the Eleventh International Mathematical Olympiad (IMO) held in Bucharest in July 1969 prompts my taking advantage of the publishing of the questions to whet the appetites of those teachers of mathematics on any academic level who might have some spunk to push for a USA Mathematical Olympiad—and even participation in the IMO's of the Eastern bloc countries. The questions of the Eleventh IMO follow [1]. The maximum number of points allowed each problem are shown. The maximum individual mark was 40, a customary arrangement, it seems.

PAPER I (4 hours) July 10, 1969.

1. Prove that there are infinitely many natural numbers a with the following property: the number $z=n^4+a$ is not prime for any natural number n.

(E. Germany, 5 points)

2. Let $a_1, a_2, \cdots, a_n$ be real constants, x a real variable, and

$$f(x) = \cos(a_1 + x) + \frac{\cos(a_2 + x)}{2} + \cdots + \frac{\cos(a_n + x)}{2^{n-1}}.$$

Prove that if $f(x_1) = f(x_2) = 0$, then $x_1 - x_2 = m\pi$, where m is an integer.

(Hungary, 7 points)

3. For each value of $k=1, 2, 3, 4, 5$ find the necessary and sufficient conditions on the number $a>0$ for there to exist a tetrahedron with k edges of length a and the remaining $6-k$ edges of length 1.

(Poland, 7 points)

PAPER II (4 hours) July 11, 1969.

4. A semicircular arc γ is drawn on AB as diameter, C is a point of γ distinct from A and B, and D is the orthogonal projection of C on AB. We consider three circles $\gamma_1, \gamma_2, \gamma_3$ which have AB as a common tangent. Of these γ_1 is the circle which is inscribed in the triangle ABC and γ_2, γ_3 are both tangential to the line-segment CD and to γ. Prove that $\gamma_1, \gamma_2, \gamma_3$ have a second tangent in common.

(Holland, 6 points)

5. Give $n>4$ points in a plane such that no three are collinear, prove that one can find at least $\binom{n-3}{2}$ convex quadrilaterals whose vertices are four of the given points.

(Mongolia, 7 points)

6. Given $x_1>0$, $x_2>0$, $x_1y_1-z_1^2>0$, and $x_2y_2-z_2^2>0$, prove that

$$\frac{8}{(x_1 + x_2)(y_1 + y_2) - (z_1 + z_2)^2} \leqq \frac{1}{x_1y_1 - z_1^2} + \frac{1}{x_2y_2 - z_2^2}.$$

Give necessary and sufficient conditions for equality.

(U.S.S.R., 8 points)

The two sessions, three problems per session arrangement with four hours allowed for each session, is again customary. The time is an important factor for the liberal amount adds to the vigor of the competition and allows for the finding of polished solutions and generalizations.

Fourteen teams of eight members each, with eight members the customary team size, took part in the Roumanian hosted Olympiad. While the IMO's are basically Eastern bloc affairs, some Western countries have been invited to participate and have participated the last few years [2]. Incidentally, invitations to Western countries and arrangements for participation of those countries has to be handled on a high governmental basis. We have never participated; and so far we have not shown any interest or courage in participating.

It seems to be customary, too, to divide the individual scores into three classes. At the tenth IMO, a score of $\geqq 39$ rated first class, a score $\geqq 34$ but <39 rated second class, and a score of $\geqq 22$ but <34 rated third class. Class scorings for the Eleventh IMO were, no doubt, similar from the total scores shown below in terms of participating countries and number of team members falling in each class.

Country	Class			Total Score
	First	Second	Third	
Hungary	1	4	2	247
German Democratic Republic	—	4	4	240
USSR	1	3	3	231
Roumania	—	4	2	219
England	1	1	1	193
Bulgaria	—	—	3	189
Yugoslavia	—	2	2	181
Czechoslovakia	—	—	3	170
Mongolia	—	—	1	120
Poland	—	1	—	119
France	—	1	—	119
Sweden	—	—	—	104
Belgium	—	—	—	57
Netherlands	—	—	—	51

Returning to the idea of a USA Mathematical Olympiad, we might consider the purpose one would serve. It certainly would represent a step higher in secondary school competition in mathematics in our country. As a subjective type examination, the type used for the individual Mathematical Olympiads that are now being held in Eastern bloc countries and England, it would provide the challenging experience needed by students in our country to think a problem through, to organize a proof, and to express that organization in the written word. It could act as the "go between" between the Annual High School Mathematics Competition and possible participation in an IMO.

Could such an Olympiad be easily arranged for and administered? By all means! Just as the British use the Annual High School Mathematics Competition as the qualifying round for the British Mathematical Olympiad, so we could use that competition for the qualifying round for an Olympiad of our own. The British select the top 60 or 70 in the Annual High School Mathematics Competition, as it pertains to their schools, for participants in their Olympiad. We could do the same. The British students sit for their Olympiad in their own schoools. Ours could do the same. Professor and Mrs. Walter Hayman who originated the British Mathematical Olympiad did assume the responsibility of grading the Olympiad papers. We should be able to find at least a committee to do that here. From the winners of their Olympiads, the British draw the members of the teams to take to the IMO's. We could do the same.

What good would come from a USA Mathematical Olympiad? It would provide a measuring stick of our provision for talented students in secondary school mathematics. It might encourage a cutting down on the amount of multiple-choice testing we use and the putting of some emphasis on subjective type testing, a type we need. It might encourage the providing of special treatment for those children who are highly talented in mathematics and need stimulation and challenge. It would be a stepping stone toward an invitation to participate in an IMO. There would be great value resulting from such participation. From among the participants of the IMO's there quite likely will come the leading mathematicians of the next generation. Participation provides the opportunity for these future leaders to meet one another and to be exposed vigorously to a foreign culture at an early age.

There exists the possibility that on a first try we would land at the bottom of the heap. Mongolia did in the 1968 IMO at Moscow but beat Poland, France, Sweden, Belgium, and the Netherlands in the 1969 IMO at Bucharest. France was at the bottom in 1967 at the Ninth IMO at Getinje, Yugoslavia, but beat Sweden, Belgium, and the Netherlands in the Eleventh IMO at Bucharest; she did not send a team to Moscow. It would be possible for us to perform such a feat. We certainly must possess here in the USA the strength of character to face defeat and the capability and courage to then plunge into systematic hard training to compete again with the desire to strive for a better showing.

A start on concern for international competition in the USA was made at the August, 1968 meeting of the MAA at the University of Wisconsin when a subcommittee on "international competitions" was appointed with the charge to report back to full committee by the next August. The spark fizzled out, however. As a member of the committee, to my knowledge, there never was a report.

The quoting of an opinion of R. C. Buck expressed on July 29, 1958 when the Wisconsin Section organized its mathematics contest for high school students, might be pertinent here: "May I add that I hope we will see a wider support for the contest program of the MAA, among those mathematicians whose central interest lies in research? This was the case in Hungary; it is presently the case in the USSR. A well-designed and exploited contest, supported by both the AMS

and the MAA, can have a far-reaching effect upon the mathematical atmosphere of our high schools; but quality must not be sacrificed to quantity, nor must awards go only to those who show their ability to handle routine calculations with fantastic speed. We are looking for creative minds, and they are difficult to trap with the best of weapons." I contend that a Mathematical Olympiad of the character of a British Mathematical Olympiad with questions the like of those of the Fourth British Mathematical Olympiad [3] would "trap creative minds."

So why not at least have a USA Mathematical Olympiad? Remember, "All's well that ends well."

References

1. F. R. Watson, Eleventh International Mathematical Olympiad, Bucharest, July 1969, Math. Gaz., 53 (1969) 395–396.

2. Nura D. Turner, Mathematics in Europe, East and West, New York State Mathematics Teachers Journal, 20 (1970) 12–17.

3. Examination for the Fourth British Mathematical Olympiad, this MONTHLY, 76 (1969) 80–81.

PROBLEMS AND SOLUTIONS

EDITED BY EMORY P. STARKE

All problems (both elementary and advanced) proposed for inclusion in this Department should be sent to E. P. Starke, 1000 Kensington Ave., Plainfield, N. J. 07060. Proposers of problems are urged to enclose any solutions or information that will assist the editors. Ordinarily, problems in well-known textbooks and results in generally accessible sources are not appropriate for this Department. No solutions (except those accompanying proposals) should be sent to Professor Starke.

ELEMENTARY PROBLEMS

Solutions of Elementary Problems should be sent to Problems Group, Mathematics Department, University of Maine, Orono, ME 04473. To facilitate their consideration, solutions of Elementary Problems in this issue should be typed (with double spacing) and should be mailed before May 31, 1971. Contributors (in the United States) who desire acknowledgement of receipt of their solutions are asked to enclose self-addressed stamped postcards.

An asterisk (*) means neither the proposer nor the editors supplied a solution.

E 2277. *Proposed by Phyllis Zweig Chinn, Towson State College, Maryland*

A *graph* is a finite collection of points, and lines between them, where each line has two distinct endpoints and no two lines have the same pair of end-

points. The *degree* (or valency) of a point is the number of edges to which it belongs. The *partition* associated with a graph is the sequence of degrees of points in the graph. A *frequency partition,* which is a partition of the order of a graph, can be formed by recording the frequency with which each degree is assumed.

Prove that for any partition of an integer p, except $p=1+1+\cdots+1$, there is a connected graph of order p having the given partition as its frequency partition.

E 2278. *Proposed by Henry Cheng, University of California, San Diego*

What is the number of shortest paths from one corner of a chessboard to the diagonally opposite corner which can be traversed by a rook in seven moves, but no fewer?

E 2279. *Proposed by Erwin Just, Bronx Community College*

It has been shown (see C. A. Grimm, *A conjecture on consecutive composite numbers,* this MONTHLY, 76 (1969), 1126–1128) that each member of the sequence of integers, $n!+2, n!+3, \cdots, n!+n$, is divisible by a prime which does not divide any other member of the sequence. Prove that for any positive integers n and k, there exists a sequence of n consecutive integers such that each member of this sequence is divisible by k distinct prime factors no one of which divides any other member of the sequence.

E 2280. *Proposed by Felix Magnotta, Washington and Jefferson College*

Solve the functional equation

$$f(x+y) = f(x-y) + y[f'(x+y) + f'(x-y)].$$

E 2281. *Proposed by Cornelius Groenewoud, Snyder, New York*

Let Q be the midpoint of the line segment PR. Construct with compass and straightedge a triangle ABC having P for orthocenter, Q for incenter, and R for centroid.

E 2282. *Proposed by W. J. Blundon, Memorial University of Newfoundland*

For any triangle (other than equilateral) with circumcenter O, incenter I, and orthocenter H, let the angles have measures $\alpha \leqq \beta \leqq \gamma$. Prove

$$(1) \qquad 1 < OH/IO < 3 \quad \text{and} \quad 0 < IH/OH < 2/3.$$

$$(2) \qquad 0 < IH/IO < 1 \quad \text{if } \beta < 60°, \qquad IH = IO \quad \text{if } \beta = 60°,$$

$$1 < IH/IO < 2 \quad \text{if } \beta > 60°,$$

and show that the constant 2 in the last inequality cannot be replaced by a smaller number.

SOLUTIONS OF ELEMENTARY PROBLEMS

Kaprekar's Constant

E 2222 [1970, 307]. *Proposed by R. M. Krause, National Science Foundation*

The following problem has been making the rounds at Berkeley, M.I.T., etc. Let N be any number between 0001 and 9998, excluding only 1111, 2222, etc. Arrange the digits in ascending order of magnitude and in descending order of magnitude and take the difference of the two resulting numbers. Call this difference $T(N)$. Show that the repeated application of the operator T converges on the number 6174. Is generalization possible?

I. *Solution by Norman Miller, Queen's University.* In the number N denote the four digits arranged in descending order of magnitude by $abcd$.

CASE I. $a \geqq b = c \geqq d$, with not all the digits equal. Then, in $T(N)$, the sum of the first and fourth digits is 9, and each of the second and third digits is 9. This gives only five combinations of four digits, each of which is found to lead, on further operations, to the number 6174.

CASE II. $a \geqq b > c \geqq d$. Then, in $T(N)$, the sum of the first and fourth digits is 10 and the sum of the second and third digits is 8. These possibilities combine to give 25 possible combinations of digits, each of which leads to the number 6174. This completes the proof.

II. *Generalization by C. W. Trigg, San Diego, California.* Kaprekar's constant, 6174, was announced [1] in 1949 and later discussed [2], [3] by him. In the last reference, he developed a three-digit predictive index composed of intra-differences. In the course of various generalizations, Jordan [4] confirmed the invariance of 6174.

In any system of notation with base r, only the ordered integers $a\ b\ c\ d$ with $a > d$ need be considered. Since the ordered integers $a\ b\ c\ d$ and $(a+p)\ (b+q)\ (c+q)\ (d+p)$ have the same $T(N)$, it is sufficient to deal with the $r(r+1)/2-1$ ordered integers $(r-1)\ (r-1)\ c\ d$ where $r-1>d$. If N is of this form, then $T(N) = (r-1+d)\ (r-2-c)\ c\ (d+1)$, except when $c=r-1$, in which case $T(N) = (r-2-d)\ (r-1)\ (r-1)\ (d+1)$.

Thus the integers to examine are reduced to those four-digit multiples of $(r-1)$ in which the sum of the extreme digits is r and the sum of the other digits is $(r-2)$, and those in which the sum of the extreme digits is $(r-1)$ and each of the other digits is $(r-1)$, with $(d+1)>0$. When these multiples are ordered, there are to be examined $(r^2+2r-4)/4$ distinct integers when r is even, and $(r^2-1)/4$ distinct integers when r is odd.

When each of these integers is operated on by T, the resulting computations may be assembled into flow charts which will converge on a *self-producing integer* (the term is Jordan's) or a *regenerative loop*. In some systems the computations group themselves into two or more flow charts. If all integers in a system lead to one terminal situation, it is called *unanimous*. From these charts the number of operations by T needed to move any given integer to the terminal situation

may be read directly. The operations required for any generator of a listed multiple will be one more than that required by the multiple.

In the decimal system, just 29 computations are necessary to construct its flow chart and hence to show that repeated applications of T converge on the unanimous 6174.

This is essentially the method employed in [5] for $r=2, 3, \cdots, 12$. Also unanimous is 3032 (base 5). Two self-producing integers occur in base 2. A self-producing integer and a regenerative loop occur in base 4. A single loop occurs in each of the bases 3, 6 and 7. In each of bases 8, 9, 11, and 12, two loops exist.

The results of applying T to two-digit, three-digit, and five-digit integers are discussed in [6], [7], [8].

References

1. D. R. Kaprekar, *Another Solitaire Game*, Scripta Mathematica, 15 (September 1949) 244–245.

2. ———, *An Interesting Property of the Number* 6174, Scripta Mathematica, 21 (December 1955) 304.

3. ———, *The New Constant* 6174, Devlali Camp, Devlali, India, 42 pages, paper bound.

4. J. H. Jordan, *Self-producing sequences of digits*, this MONTHLY, 71 (1964) 61–64.

5. C. W. Trigg, *Predictive indices for Kaprekar's routine*, Journal of Recreational Mathematics, scheduled for publication, October, 1970.

6. ———, *Kaprekar's routine with two-digit integers*, submitted to School Science and Mathematics.

7. ———, *All integers lead to* $\cdots$, submitted to the Mathematics Teacher.

8. ———, *Kaprekar's routine with five-digit integers*, submitted to the Mathematics Magazine.

Also solved by Anders Bager (Denmark), W. E. Buker, Leon Gerber, Michael Goldberg, Bill Kaemfer & E. F. Schmeichel, Barbara A. Keller, H. L. Nelson, Dan Nemzer, E. T. Ordman, Problems Group of Concordia College, K. C. Robinson (Australia), Marilyn Rodeen, L. E. Shader, M. F. Stilwell, W. G. Wild, E. T. Wong, and the proposer.

Editorial Note. Ordman reports that the problem is also the subject of V. A. Golubev, *Sur un curieux résultat arithmétique*, Mathesis, 66 (1957) 25–28.

The Monge Shuffle

E 2223 [1970, 307]. *Proposed by M. Slater, University of Bristol, England*

A pack of cards is shuffled by starting with the first card and then placing successive cards alternately above and below the growing discard pile. After repeated shuffles of this type, will the pack first return to its original order when the original top card first returns to its top position?

Solution by Andrew Odlyzko, Pasadena, California. A shuffle is a permutation which takes the card in position n to position $26+\frac{1}{2}(n+1)$ if n is odd and to position $27-\frac{1}{2}n$ if n is even. Using this fact, we easily find the permutation to be

$$(1\ 27\ 40\ 7\ 30\ 12\ 21\ 37\ 45\ 49\ 51\ 52)(2\ 26\ 14\ 20\ 17\ 35\ 44\ 5\ 29\ 41\ 47\ 50)$$

$$(3\ 28\ 13\ 33\ 43\ 48)(4\ 25\ 39\ 46)(6\ 24\ 15\ 34\ 10\ 22\ 16\ 19\ 36\ 9\ 31\ 42)$$

$$(8\ 23\ 38)(11\ 32)(18).$$

Since 1 returns to its position after 12 shuffles, and all cycle lengths divide 12, the pack *will* first return to its original order when the original top card first returns to its top position.

Essentially the same solution was submitted by H. T. Fitzpatrick, S. I. Gendler, Michael Goldberg, Marjorie K. Gregg, F. T. Howard, Thomas Hughes, R. L. Jow, J. V. Michalowicz, J. Pfaendtner, F. W. Saunders, R. W. Sielaff, Tony Waters, Charles Wexler, and W. G. Wild.

R. D. Gee, F. W. Humburg, and M. F. Stilwell wrote computer programs to solve the problem. R. J. McEliece & E. R. Rodemich, and K. M. Wilke employed an alternate solution.

Other solvers, W. E. Buker, Mannis Charosh, and Eric Rosenthal, point out that the problem was worked out in 1773 by Gaspard Monge and is discussed in W. W. Rouse Ball, *Mathematical Recreations and Essays*, Macmillan, 1962, pp. 310–311. Nancy Plunkett & E. F. Schmeichel, and Henry Ricardo refer to Uspensky and Heaslet, *Elementary Number Theory*, McGraw-Hill (1939), while Ricardo, Simeon Reich (Israel), and E. F. Wilde himself refer to Wilde and Tomandl, *On shuffling cards*, Mathematics Magazine, 42 (1969) 139–142.

A Sequence of Real Numbers

E 2224 [1970, 308]. *Proposed by D. P. Giesy, University of Southern California*

Suppose $\{a_n\}$ is a sequence of nonnegative real numbers such that $\limsup_{n\to\infty}(a_1+\cdots+a_n)/n<\infty$ and $\lim_{n\to\infty}a_n/n=0$. Does it necessarily follow that $\lim_{n\to\infty}(a_1^2+\cdots+a_n^2)/n^2=0$?

Solution by E. F. Schmeichel, Itasca, Illinois. The answer is yes. If a_n is bounded, this is trivial. Otherwise, choose a subsequence $\{a_{n_k}\}$ of a_n as follows: let n_1 be the smallest integer for which $a_{n_1}>a_1$, and in general let n_k be the smallest integer greater than n_{k-1} for which $a_{n_k}>a_{n_{k-1}}$. If $n_k\leqq n<n_{k+1}$, we find

$$(1)\qquad 0\leqq\frac{a_1^2+\cdots+a_n^2}{n^2}<\left(\frac{a_1+\cdots+a_n}{n}\right)\frac{a_{n_k}}{n_k}.$$

With $(a_1+\cdots+a_n)/n$ bounded and $\lim_{n\to\infty}a_n/n=0$, the right hand term in (1) converges to 0, and so the middle term of (1) also converges to 0.

Also solved by Annie L. Alexander, K. F. Anderson, M. T. Bird, W. M. Causey, R. J. Driscoll, Leon Gerber, Ellen Hertz, Ed Jones & Robert Young, David Kelly, B. G. Klein, J. R. Kuttler, Harry Lass, Beatrize Margolis (Argentina), J. C. Molluzzo, Walter Noll, Andrew Odlyzko, Steve Rohde, Sid Spital, St. Olaf College Students, G. E. Volland, and the proposer.

Sequential Integers

E 2225 [1970, 308]. *Proposed by Diane Comer and J. J. Tourneau, Fisk University*

Show that any positive integer S can be written in exactly k different ways as the sum of two or more consecutive positive integers (in increasing order), where k is the number of positive odd divisors of S greater than 1.

Solution by Anders Bager, Hjørring, Denmark. It follows from

$$x+(x+1)+\cdots+(x+y)=\tfrac{1}{2}(y+1)(2x+y)$$

that we must seek positive integers x, y such that

(1) $$\tfrac{1}{2}(y+1)(2x+y)=S.$$

We distinguish two cases.

(I) y even, $y=2z$ where z is a positive integer. (1) then becomes

(2) $$(2z+1)(x+z)=S.$$

Here $u=2z+1$ is an odd divisor >1 of S. Conversely, suppose that $S=uv$, $u>1$ and odd. We put $2z+1=u$, $x+z=v$, solve, and get $z=\frac{1}{2}(u-1)$, $x=\frac{1}{2}(2v-u+1)$. Clearly, (x, z) is an acceptable solution to (2) if and only if

(3) $$2v-u\geqq 1.$$

(II) y odd, $y=2z-1$ with z a positive integer. (1) now becomes

(4) $$z(2x+2z-1)=S.$$

Here $u=2x+2z-1>1$ is an odd divisor of S. Conversely, suppose that $S=uv$, $u>1$ and odd. We put $2x+2z-1=u$, $z=v$, solve, and get $x=\frac{1}{2}(u-2v+1)$. Clearly, (x, z) is an acceptable solution to (4) if and only if

(5) $$2v-u\leqq -1.$$

Now, if $u>1$ is an odd divisor of $S=uv$, one and only one of the conditions (3) and (5) is satisfied. Hence u provides a unique solution (x, y) to (1).

Also solved by Walter Bluger, Mannis Charosh, Josef Daneš (Czechoslovakia), G. C. Dodds, Kathleen Ann Drude, S. I. Gendler, M. G. Greening (Australia), J. J. Herbold, Dean Hickerson, Edwin Hoefer & Samuel Lawn, F. T. Howard, F. W. Humburg, Yul J. Inn, David Kelly, R. M. King, B. G. Klein, Harry Lass, Arthur Marshall, B. R. Myers, Robert Patenaude, C. B. A. Peck, Simeon Reich (Israel), E. F. Schmeichel, Michael Shimshoni (Israel), Stephen Spindler, Charles Wexler, W. K. Wild, Gerald Wildenberg, and the proposer.

Partial solutions were submitted by R. A. Gibbs, Michael Goldberg, M. S. Klamkin, V. J. Motto, Marilyn Rodeen, Jim Tatersall, and K. M. Wilke.

Intersecting Altitudes of a Tetrahedron

E 2226 [1970, 308]. *Proposed by M. S. Klamkin, Ford Scientific Laboratory*

If one altitude of a tetrahedron intersects two other altitudes, then all four altitudes are concurrent.

I. *Solution by Judith R. Gumerman, West Chester (Pa.) State College.* Let the altitudes of the tetrahedron be a, b, c, and d, and let a intersect b and c. Project the entire figure parallel to a (onto the face perpendicular to a). Then a projects onto a single point, and b, c, and d project onto three altitudes of the base triangle (the face perpendicular to a). The projection of a must intersect the projections of b and c. Since the projection of a is a point, it must be the point of intersection of the projections of b and c. Since the three altitudes of a triangle are concurrent, the projection of d also passes through the projection of a, so a intersects d as well as b and c. Next, project the entire figure parallel to b. Since a intersects

both c and d, and since their three projections are concurrent (they are the altitudes of a triangle), the altitudes a, c, and d are concurrent. Similarly, by projecting parallel to c, we find that a, b, and d are concurrent. Thus, all four altitudes meet in a point.

II. *Solution by Simeon Reich, Israel Institute of Technology.* Let $ABCD$ be the given tetrahedron, and let h_A intersect h_B and h_C. Then AB is perpendicular to CD and AC is perpendicular to BD (Nathan Altshiller-Court, *Modern Pure Solid Geometry*, 2nd Ed., §204). This can be expressed by

$$AB\cdot(AD - AC) = 0, \qquad AC\cdot(AD - AB) = 0.$$

Hence $\mathbf{AD}\cdot(\mathbf{AB}-\mathbf{AC})=0$. That is, AD is perpendicular to BC. It follows that the altitudes are concurrent (*loc. cit.* §208, §212).

Also solved by Anders Bager (Denmark), Leon Bankoff, D. J. Bordelon, Jordi Dou (Spain), F. M. Eccles, Leon Gerber, Michael Goldberg, M. G. Greening (Australia), Kit Hanes, D. G. Kabe, David Kelly, František Kuřina (Czechoslovakia), Harry Lass, Daniel Pedoe, E. F. Schmeichel, P. D. Thomas. E. W. Trost (Switzerland), Charles Wexler, W. G. Wild, and the proposer.

Divisors of Binomial Coefficients

E 2227 [1970, 308]. *Proposed by N. S. Mendelsohn, University of Manitoba*

Find the greatest common divisor of

$$\binom{2n}{1}, \binom{2n}{3}, \binom{2n}{5}, \cdots, \binom{2n}{2n-1}.$$

Solution by St. Olaf College Students. From the familiar relation

$$\binom{2n}{1}+\binom{2n}{3}+\binom{2n}{5}+\cdots+\binom{2n}{2n-1}=2^{2n-1}$$

it follows that their common divisor must be of the form 2^p. If $n=2^kq$, where q is an odd integer, then from $\binom{2n}{1}=2^{k+1}q$ it follows that a common divisor of these coefficients cannot be larger than 2^{k+1}.

To show that 2^{k+1} divides all of them we write

$$\binom{2^{k+1}q}{p}=\frac{2^{k+1}q}{p}\binom{2^{k+1}q-1}{p-1}.$$

Since binomial coefficients are integers and p is odd, it follows that

$$\binom{2^{k+1}q}{p}=2^{k+1}M,$$

where M is an integer and $p=1, 3, \cdots, 2n-1$.

This proves that 2^{k+1} is the greatest common divisor of the given set.

Also solved by Joe Albree, Einar Andresen (Norway), J. C. Binz (Switzerland), D. M. Bloom, T. R. Butts, L. Carlitz, D. M. Cohen, Josef Daneš (Czechoslovakia), G. C. Dodds, J. M. Gandhi,

M. F. Gillis, R. E. Giudici (Chile), Michael Goldberg, M. G. Greening (Australia), H. N. Gupta, F. T. Howard, J. R. Kuttler, Harry Lass, Graham Lord, H. G. ter Morsche (Netherlands), R. C. Mullin, Andrew Odlyzko, D. P. Pazel, Simeon Reich (Israel), Kenneth Rosen, E. F. Schmeichel, K. W. Schmidt, Karen J. Schroeder, R. E. Shafer, Michael Shimshoni (Israel), G. J. Simmons, W. W. Tom, E. W. Trost (Switzerland), J. R. Ventura, Jr., Charles Wexler, David Zeitlin, and the proposer.

Comment by Anders Bager, Denmark

Dear Problems Group,

I once asked you the stupid question if you were an Abelian group. Clearly this is so as you have 9 elements. [There were only nine members to the Group at the time of his writing. Ed.] I am of the opinion that you must be an elementary Abelian Group (because you are an ELEMENTARY PROBLEMS GROUP, of course) and hence cannot be cyclic. Although it is customary to write the unit first, you certainly have put it last; I have identified it as the *sole* element of junior order 1. It must be some sort of feat to identify the identity!

A. B.

ADVANCED PROBLEMS

All solutions of Advanced Problems should be sent to J. Barlaz, Rutgers—The State University, New Brunswick, N. J. 08903. Solutions of Advanced Problems in this issue should be typed (with double spacing) on separate, signed sheets and should be mailed before May 31, 1971. Contributors (in the United States) who desire acknowledgement of receipt of their solutions are asked to enclose self-addressed, stamped postcards.

An asterisk (*) means neither the proposer nor the editors supplied a solution.

5754 [1970, 890]. **Correction.** *Proposed by M. S. Klamkin, Ford Scientific Laboratory*

If $L(a, b)$ denotes the perimeter of an ellipse with semi-axes a and b ($a \geqq b$), show that

$$L^2(a, c) - 16a^2 \geqq L^2(b, c) - 16b^2.$$

5777.* *Proposed by H. W. Gould, West Virginia University*

It is not difficult to prove that m/d is a factor of the binomial coefficient $\binom{m}{n}$, where $d=(m, n)$ is the greatest common divisor of m and n. Also $(m-n+1)/(m+1, n)$ is a factor of $\binom{m}{n}$. According to Dickson, *History of the Theory of Numbers*, (Vol. 1, p. 272) these results trace back to Hermite in 1889.

Can one find a general result of the form

$$\frac{am+bn+c}{(rm+s, un+v)} \Bigg| \binom{m}{n}?$$

5778. *Proposed by L. W. Shapiro, Howard University*

Find the smallest group of finite order with no subgroup of prime index.

5779. *Proposed by G. J. Janusz, University of Illinois*

Let p be an odd prime and $F(X)$ the polynomial with rational coefficients such that $F(2\cos\theta) = 2\cos p\theta$ for all real θ. Let m and n be nonzero integers both relatively prime to p such that $|pm| < n$. Set $f(X) = F(X) - 2\,pm/n$. Prove

(1) $f(X)$ is irreducible over the rationals,

(2) the roots of $f(X)$ are all real,

(3) the Galois group of $f(X)$ over the rationals is solvable with order dividing $2p(p-1)$.

5780. *Proposed by W. R. Emerson, New York University*

For which algebraic number fields $F([F:Q] < \infty)$ is the following valid? A polynomial $P \in \theta[x]$ is reducible over $F[x]$ if and only if it is reducible over $\theta[x]$, where θ is the ring of integers of F.

5781. *Proposed by P. R. Chernoff, University of California, Berkeley*

Generalizing a well-known result of Liouville, prove that a harmonic function $u(x)$ of polynomial growth on R^n must be a polynomial.

5782. *Proposed by Jiang Luh, North Carolina State University*

Let G be a torsion group and H be a subgroup of G of index m (finite). Show that if all prime factors of the orders of elements in H are $\geqq m$ then H is a normal subgroup of G.

SOLUTIONS OF ADVANCED PROBLEMS

Expansion of a Poisson Kernel

5710 [1970, 85]. *Proposed by R. E. Shafer, Lawrence Radiation Laboratory, California Institute of Technology*

It is well known that

$$[R^2 - 2Rr\cos\theta + r^2]^{-\nu} = \sum_{n=0}^{\infty} \frac{r^n}{R^{n+2\nu}} C_n^{\nu}(\cos\theta),$$

$|r| < |R|$, $\mathrm{Re}(\nu) > -1$, $\mathrm{Re}(\nu) \neq 0$. Find the set of functions $F_n(r, R)$ independent of θ such that

$$[R^2 - 2Rr\cos\theta + r^2]^{-\mu} = \sum_{n=0}^{\infty} F_n(r, R) C_n^{\nu}(\cos\theta), \qquad \mathrm{Re}(\mu) > -1.$$

Solution by O. P. Lossers, Technological University, Eindhoven, Netherlands. According to an old result of Gegenbauer (cf. R. Askey, Proc. Amer. Math. Soc., 16 (1965) 1191–1194) one has

$$C_k^{\mu}(x) = \sum_{n=0}^{k} \beta_{n,k}^{\nu,\mu} C_n^{\nu}(x),$$

where

$$\beta_{n,k}^{\nu,\mu}=\frac{(n+\nu)\Gamma(\nu)\Gamma((k-n)/2+\mu-\nu)\Gamma(\mu+(k+n)/2)}{\Gamma(\mu)[(k-n)/2]!\Gamma(\mu-\nu)\Gamma(\nu+1+(k+n)/2)}$$

if $k-n$ is even, $k\geqq n$, and otherwise $\beta_{n,k}^{\nu,\mu}=0$. Using this result one has

$$\begin{aligned}[R^2-2Rr\cos\theta+r^2]^{-\mu}&=\sum_{k=0}^{\infty}\frac{r^k}{R^{k+2\mu}}C_k^{\mu}(\cos\theta)\\&=\sum_{k=0}^{\infty}\frac{r^k}{R^{k+2\mu}}\sum_{n=0}^{k}\beta_{n,k}^{\nu,\mu}C_n^{\nu}(\cos\theta)=\sum_{n=0}^{\infty}F_n(r,R)C_n^{\nu}(\cos\theta),\end{aligned}$$

where

$$\begin{aligned}F_n(r,R)&=\sum_{k=n}^{\infty}\beta_{n,k}^{\nu,\mu}\frac{r^k}{R^{k+2\mu}}=\sum_{k=0}^{\infty}\beta_{n,n}^{\nu,\mu}+2k\frac{r^{n+2k}}{R^{n+2k+2\mu}}\\&=\sum_{k=0}^{\infty}\frac{(n+\nu)\Gamma(\nu)\Gamma(k+\mu-\nu)\Gamma(\mu+n+k)}{\Gamma(\mu)k!\Gamma(\mu-\nu)\Gamma(\nu+1+n+k)}\cdot\frac{r^{n+2k}}{R^{n+2k+2\mu}}\\&=\frac{(n+\nu)\Gamma(\nu)\Gamma(n+\mu)}{\Gamma(\mu)\Gamma(\nu+1+n)}\frac{r^n}{R^{n+2\mu}}\sum_{k=0}^{\infty}\frac{(\mu-\nu)_k(\mu+n)_k}{k!(\nu+1+n)_k}\left(\frac{r^2}{R^2}\right)^k\\&=\frac{(\mu)_n}{(\nu)_n}\frac{r^n}{R^{n+2\mu}}F\left(\mu-\nu,\mu+n;\nu+1+n;\frac{r^2}{R^2}\right).\end{aligned}$$

In the above, F stands for the hypergeometric function, and $(\alpha)_n$ is Porthammer's symbol which may be defined by

$$(\alpha)_n=\frac{\Gamma(\alpha+n)}{\Gamma(\alpha)}=(\alpha+n-1)(\alpha+n-2)\cdots(\alpha).$$

Also solved by the proposer.

N. Wallach and K. Johnson observe that the formula of the problem is also contained in a paper by Helgason, *Duality for symmetric spaces with applications to group representations*, Formula 29 in Chapter 4, to appear in Advances in Mathematics. For his proof, Helgason refers to a formula on p. 81 in Erdelyi's *Higher Transcendental Functions*, New York, 1953.

A System of Congruences

5711 [1970, 85]. *Proposed by Dan Marcus, York University, Toronto*

Let $A=(a_{m,n})_{m,n=1}^{\infty}$ be an infinite matrix of nonzero integers such that for each m, the set of prime divisors of numbers in the mth row is finite. Prove that the system of congruences $x_{m+n}\equiv a_{m,n} \pmod{x_m}$ is solvable in primes.

Solution by R. B. Eggleton, University of Melbourne, Australia.

Let P_m be the set of prime divisors of numbers in the mth row of A. Let p_1 be any prime not in P_1. For any $k\geqq 2$, let $H(k)$ denote the hypothesis: *There exist distinct primes p_m $(1\leqq m\leqq k)$ such that $p_m\notin P_m$ and*

$$p_m\equiv a_{n,m-n} \pmod{p_n}\ \textit{for}\ 1\leqq n\leqq m-1\quad\textit{and}\quad 2\leqq m\leqq k.$$

Since $p_1 \notin P_1$, evidently $a_{1,1}$ and p_1 are relatively prime. Dirichlet's theorem on primes in an arithmetic progression thus ensures the existence of infinitely many primes in $\{a_{1,1}+rp_1\}_{r=0}^{\infty}$. Since P_2 is finite, the progression contains primes not in P_2; choose p_2 to be any one of them. Then $H(2)$ is seen to be valid.

Suppose $H(k)$ is valid for some particular k. Then in the congruence system $x \equiv a_{n,k-n+1} \pmod{p_n}$, $(1 \leqq n \leqq k)$, the moduli are k distinct primes, so the Chinese Remainder Theorem ensures a single congruence $x \equiv b \pmod{M}$, where $M = \prod_{n=1}^{k} p_n$. Moreover $p_n \notin P_n$ $(1 \leqq n \leqq k)$ so *a fortiori* p_n is not a divisor of $a_{n,k-n+1}$, whence $p_n \nmid x$. Thus the solution is not divisible by any prime divisor of M, so $(b, M) = 1$. By Dirichlet's theorem, there are infinitely many primes in $\{b+rM\}_{r=0}^{\infty}$, so we can choose p_{k+1} to be any one of them not in the (finite) set P_{k+1}. This implies $H(k+1)$.

Induction on k therefore establishes the validity of $H(k)$ for every $k \geqq 2$, which is equivalent to the solution, in primes, of the system $x_{m+n} \equiv a_{m,n} \pmod{x_m}$.

Also solved by Dean Hickerson, E. F. Schmeichel, E. W. Trost (Switzerland), Charles Vanden Eynden, and the proposer.

The proposer notes that the set of solutions of the system actually has cardinality c since infinitely many primes are available at each step of the inductive process.

Chromatic Number of a Simple Graph

5713 [1970, 85]. *Proposed by D. P. Geller, University of Michigan*

For any graph G with p points, q lines, and chromatic number χ, show

$$\chi \geqq \frac{p^2}{p^2 - 2q}.$$

Solution by E. F. Schmeichel, Itasca, Illinois. The desired inequality is equivalent to $q \leqq (p^2/2)(1-1/\chi)$, and we can reformulate the problem as follows: In any graph with p vertices and chromatic number χ, the number of edges does not exceed $(p^2/2)(1-1/\chi)$. Noting that every graph with chromatic number χ is a χ-partite graph, it suffices to show that the number of edges in a χ-partite graph of p vertices does not exceed $(p^2/2)(1-1/\chi)$.

Now in any χ-partite graph, the vertices can be partitioned into χ disjoint sets $\{V_1, V_2, \cdots, V_\chi\}$ in such a way that no edge joins two vertices of the same set. If n_i denotes the number of vertices in V_i, it follows that the number of edges in the graph cannot exceed $\sum_{i<j} n_i n_j$, with $n_1+n_2+\cdots+n_\chi = p$. However, we have

$$\sum_{i<j} n_i n_j \leqq \binom{\chi}{2} (p^2/\chi^2) = (p^2/2)(1 - 1/\chi),$$

with equality only if χ divides p and $n_i = p/\chi$ for all i. Thus a χ-partite graph on p vertices has at most $(p^2/2)(1-1/\chi)$ edges, as required.

Also solved by Václav Chvátal, D. M. Cvetković (Yugoslavia), R. B. Eggleton (Australia), Basil R. Myers, B. R. Myers & R. Liu, Slobodan Simić (Yugoslavia), and the proposer.

Chvátal finds the problem in Harary, *Graph Theory*, Ex. 12.29. He and the proposer also refer to the stronger result of A. Ershov and G. Kozhukhin:

$$\chi \geq \frac{p}{[t]}\left(1 - \frac{t-[t]}{1+[t]}\right),$$

where $t=p-2q/p$ (Soviet Mathematics Doklady, 3(1962) 50–53). Eggleston, with his solution, offers the upper bound

$$\chi \leq \tfrac{1}{2}(1+\sqrt{8q+1}).$$

Partitioning the Hypotenuse

5714 [1970, 197]. *Proposed by Marlow Sholander, Case Western Reserve University*

Let O be the origin, let $ABCD$ be a line segment with endpoints on the coordinate axes, and let OBC be equilateral. Let $r=AB$, $s=BC$, and $t=CD$ be positive integers. A triple r, s, t has a dual t, s, r and a sum $r+s+t$. It is called primitive if $(r, s, t)=1$. Two triples are called dependent if one is proportional to the other or to its dual. Pairs of independent primitive triples are called pips. Two such triples may have an element in common or they may have a common sum.

(i) What primes are found as elements shared by pips?

(ii) What is the minimum sum shared by pips?

Solution (abridged) by Charles Vanden Eynden, Illinois State University at Normal. Let $A=(0, a)$ and $D=(d, 0)$, and set $\Sigma=r+s+t$. Then

$$B = (rd/\Sigma, a(s+t)/\Sigma), \qquad C = (d(r+s)/\Sigma, at/\Sigma).$$

Since $|OB| = |OC| = s$ and $|AD| = \Sigma$, we have

$$a^2 + d^2 = \Sigma^2, \qquad r^2d^2 + (s+t)^2a^2 = s^2\Sigma^2,$$
$$(r+s)^2d^2 + t^2a^2 = s^2\Sigma^2.$$

Elimination of a and d from the above three equations gives

$$s(s-r-t) = 2rt. \tag{1}$$

Conversely, if r, s and t are positive integers satisfying (1), then setting

$$a = \left(\frac{\Sigma(s^2-r^2)}{\Sigma-2r}\right)^{1/2}, \qquad d = \left(\frac{t\Sigma(2s+t)}{\Sigma-2r}\right)^{1/2},$$

and defining A, B, C, and D as in the first paragraph, produces a line segment $ABCD$ with the geometry specified in the problem. Thus it is sufficient to determine the solutions of (1).

Assume $\langle r, s, t\rangle$ is a solution of (1) and let x, y be relatively prime integers such that $sx=ty$. Then (1) becomes $y(s-r-t)=2rx$. These last equations give at once

$$r:s:t = y(y-x):y(2x+y):x(2x+y). \tag{2}$$

For this triple to be primitive it may be necessary to remove a highest common factor $(=\rho)$. Suppose this has been done.

If p is any odd prime, the choice $x=p-1$, $y=p$ yields

$$\langle p, p(3p-2), (3p-2)(p-1)\rangle,$$

while $x=2p-1$, $y=2p$ yields $\langle p, 2p(3p-1), (2p-1)(3p-1)\rangle$. These are primitive triples and thus the answer to (i) is "all primes."

To answer (ii) we note that the triples $\langle 110, 143, 13\rangle$ and $\langle 45, 153, 68\rangle$ both have $\Sigma=266$. From (2) we have $\Sigma=2(x^2+xy+y^2)$ or this reduced by ρ. It follows that $\Sigma\to\infty$ as $x\to\infty$. Thus $\Sigma<266$ for only finitely many triples. Direct calculations produce no smaller shared Σ.

Also solved by Einar Anderson (Norway), Walter Bluger, W. J. Blundon, M. S. Demos, J. G. Mauldon, and the proposer.

REVIEWS

EDITED BY J. ARTHUR SEEBACH, JR., AND LYNN A. STEEN

with the assistance of the mathematics departments of St. Olaf and Carleton Colleges.

COLLABORATING EDITOR FOR FILMS: SEYMOUR SCHUSTER

Printed materials for review should be sent to: Book Review Editor, American Mathematical Monthly, St. Olaf College, Northfield, MN 55057. Films and correspondence relating to films should be sent to Seymour Schuster, Carleton College, Northfield, MN 55057.

All unsigned material is written by one of the editors. A boldface capital **C** *in the margin indicates that a review is based in part on classroom use. Professors willing to write such a review should inform the editor in order to avoid duplication.*

C *Introduction to the Calculus of Variations.* By Hans Sagan. McGraw-Hill, New York, 1969. xvi+449 pp. $14.95. (Telegraphic Review, August–September 1970.)

The last few years have finally witnessed the appearance of the long-awaited books on the calculus of variations by M. R. Hestenes, L. C. Young, and G. M. Ewing. H. Sagan's book is another welcome addition to this area. There is a great deal of individuality in each of these books which makes comparison among them difficult. On the other hand, the books complement each other in many ways, providing students with a fine balance between classical theory, recent developments, and personal reflections on a subject which has greatly influenced the development of modern mathematics.

There are several features which make the book under review eminently suitable as a text for an introductory course: the style is pleasant; the prerequisites are kept to a minimum (advanced calculus and intermediate differential equations); and the pace of the development is appropriate for most students at the senior or first year graduate level.

The book is designed to lay "a foundation for an understanding of the prob-

lems, methods and techniques of the calculus of variations and to prepare the reader for the study of modern optimal control theory. The treatment is limited to an extensive coverage of single integral problems in one or more unknown functions." The exposition is predominantly classical and only Riemann integrals are employed. However, the author uses elementary concepts of normed spaces on several occasions as a pedagogical aid, enabling him, for instance, to give a novel treatment of transversality conditions and constrained variational problems. Pontryagin's minimum principle is developed only to the extent possible within the framework of the Hamilton-Jacobi theory.

Direct and numerical methods are not treated. Multiple integral problems and their associated natural boundary conditions are treated only peripherally. This limits the usefulness of the book for students whose primary interest in the calculus of variations lies in its applications to areas (such as continuum mechanics) where field problems are paramount.

This reviewer has used most of chapters 1, 2, 5, and 6, (supplemented with material on multiple integral problems and eigenvalue problems), in a one-quarter graduate course for students primarily interested in engineering and physics, and has also used the book for part of the content of a basic graduate sequence in applied mathematics (based on the theory of linear spaces and operators). Most of the students found the book quite satisfactory except for minor misprints. The book has over 400 exercises, mostly of a simple nature.

For his intended goal and readers, the author has succeeded in writing a book which should find wide acceptance as a text and reference.

M. Z. NASHED, Georgia Institute of Technology

Historical Topics for the Mathematics Classroom. Thirty-First Yearbook. National Council of Teachers of Mathematics, Washington, D. C., 1969. 542 pp. $7.50. (Telegraphic Review, January 1970.)

This book is well constructed in terms of achieving its purpose of "providing a yearbook on the use of the history of mathematics in the teaching of mathematics."

The volume consists of eight essays on broad topics (called "overviews") and each, except the first and last, is followed by a series of short "capsules" on related topics. There are 114 capsules in all, by 84 authors. Each capsule contains a short bibliography. This structure allows the reader to use the material independently of its order of appearance, and according to his own needs and interests. The index is very good.

This book should not be used as a text in the history of mathematics, nor was it designed for such a purpose. In the opening article (The History of Mathematics as a Teaching Tool) Philip S. Jones states: " . . . properly used, a sense of the history of mathematics and its uses, is a significant tool in the hands of a teacher who teaches 'why' · · · (it) will not function as a teaching tool unless the users (1) see significant purposes to be achieved by its introduction and (2) plan thoughtfully for its use to achieve such purposes." With that role in mind, this volume can be invaluable for college teachers as well as the

elementary and secondary teachers to whom it is addressed. Though much of the treatment is thin, and there is some overlapping (which was unavoidable the way the volume was constructed), the weaknesses can be turned to pedagogical advantage. Recognizing that "the history of mathematics is not a static subject" can lead to an open-ended approach which may spur both student and teacher to investigate various topics more carefully and thoroughly.

With the exception of the closing overview: "Development of Modern Mathematics" by R. L. Wilder, there is little material on the mathematics since 1800. This is explained in the preface as being due to "lack of space and want of contributors." The latter reason indicates a glaring and regrettable deficiency in the literature as well as a lack of people interested in making such contributions.

H. S. TROPP, University of Toronto

Introduction to IBM/360 Assembler Language. By James Rosenberg. Holden-Day, San Francisco, 1970. 138 pp. $3.95 (paper). (Telegraphic Review, October 1970.)

The book appears to be designed to provide some insight into assembler language and a few related concepts in the use of a computer to the reader who has had substantial experience with a programming language such as Fortran. The major parts of the book are the technical description of the storage structure and machine instructions, a brief introduction to assembler language, programming techniques particularly related to assembler language, the interrupt concept, and input/output.

The author achieves his purpose rather well, and he appears to be correct in his claim in the preface: "This book will not, by itself, produce assembler language programmers." The author has dealt with the excessive details of the machine and assembler language cleverly and carefully, although at times the machine instruction formats are difficult to sort out. There are several appropriate examples in the chapters on programming techniques. These are described clearly and completely, but a few more would have been desirable. The few problems are not particularly impressive or useful.

This is an excellent book for that inquiring student who has had experience with a programming language and wants to look beyond the language. It should be satisfactory either for self-study or a formal class. Most of all, it should be readily available wherever a 360 system is used.

E. R. MULLINS, JR., Swarthmore College

Introduction to Combinatorial Mathematics. By C. L. Liu. McGraw-Hill, New York, 1968. x+393 pp. $13.50. (Telegraphic Review, March 1969.)

The fourteen chapters of this book fall into four classes. The first five chapters deal with combinatorial analysis in the traditional sense—permutations and combinations (or arrangements and selections, in the author's terminology), generating functions, recurrence relations, inclusion and exclusion, and Polya's theorem. Various topics from graph theory, including independent sets, chrom-

atic polynomials, vector spaces associated with a graph, and Kuratowski's theorem, are presented in the next four chapters. Optimization techniques pertaining to transport networks, matching theory, and linear and dynamic programming are considered in the next four chapters (Menger's theorem is not mentioned although one might have expected to have found it here). The final chapter is devoted to block designs.

The book is not intended to be an advanced treatise on any of these topics (Riordan, Ryser, and Hall, for example, each treat various subjects in greater detail in their books on combinatorics); rather, it is a broad elementary survey of the field. Each chapter ends with a selected bibliography and a collection of problems (an answer key is available). The book would seem quite suitable for a first course in combinatorics for advanced undergraduate or beginning graduate students.

If one wanted to quibble one might wonder how helpful it is to suggest on p. 26 that we needn't worry about the convergence of a generating function $F(x)$ because we can always set x equal to zero. Also, it's not made clear why over a page is taken on pp. 375–376 to show that $JQ=kJ$, where Q is the incidence matrix of (v, k, λ)-configuration, when this follows immediately from the fact that Q has k 1's in each column according to the definition given. Most sentences that end with an algebraic expression are not punctuated; some may think this sets a poor example for students.

J. W. MOON, University of Alberta

C Professor George Polya and I used this book in our Combinatorial Analysis course. We found the book very satisfactory because it covered the usual topics such as permutations and combinations, generating functions, counting theorems, but also the theory of graphs, introduction to linear programming, and network flow theory. In other words, the book covered problems which are concerned with counting the number of different combinations and also problems concerned with selecting that combination which is optimal in some sense. The course was very popular and I believe that it will become an important course in due time.

GEORGE DANTZIG, Stanford University

Linear Algebra and Geometry. By James A. Murtha and Earl R. Willard. Holt, Rinehart, and Winston, New York, 1969. 245 pp. \$7.95. (Telegraphic Review, November 1969.)

As the title indicates this text is a blend of elementary linear algebra and its connections with affine and projective geometry. Of the four chapters two are devoted to linear algebra and one to each of affine and projective geometry. The authors mention in the preface that the material was originally written with high school teachers in mind and this seems to be a fair assessment of the level of the text.

Chapter 1 (Finite dimensional vector spaces) develops the usual elementary facts about vector spaces over division rings and linear transformations as well

as some simple group theory. Although matrices are introduced and discussed at some length the presentation is essentially coordinate-free. Canonical forms for linear transformations over fields are not discussed. Chapter 2 (Affine Geometry) uses the material in Chapter 1 to develop affine geometry over a division ring. Desargues' theorem is proven and discussed. The structure of the affine group is investigated and related to the general linear group. Chapter 3 (Multilinear Algebra) deals with bilinear and quadratic forms over a field and proves the existence and uniqueness of the determinant function. The Gram-Schmidt, real spectral, and Cayley-Hamilton theorems are proven. Chapter 4 (Projective Geometry) develops some of the elementary theory of projective space. Most of the material deals with Desarguian spaces. Topics covered include lattices and duality, collineations, co-ordinatization, the fundamental theorem of projective geometry, cross ratio, and conics. Included are proofs of the fact that a Desarguian space of dimension more than two is Pappian if and only if the coordinate ring is a field, and the theorems of Steiner and Pascal.

In general the book is well written and most of the topics well motivated. The exercises are on the whole good and stress the conceptual rather than computational aspects of the subject. Some of the exercises are used later on in the body of the text.

PETER BOTTA, University of Toronto

The History of the Abacus. By J. N. Pullan. Praeger, New York, 1969. xiii+127 pp. $4.95. (Telegraphic Review, October 1969.)

The favorable "telegraphic review" (October, 1969) is entirely justified provided the title is read as printed. The book is not a *mathematical* history of the abacus. The author of this nice booklet does not claim to be a mathematician. He is a retired English inspector of schools, and is chairman of an English Archaeological Society. It is distinctly a book written by an Englishman for Englishmen, with phrases like "Roman customs in our own country." All source references, except for the centuries old classical ones, are to English (and in a very few cases to American) authors.

People had to deal with numbers now written with the Hindu-Arabic positional zero long before this system was introduced. In decimal systems without zero there was no way to distinguish in writing between 3200, 3020, 3002, 320, 302 and 32. Any abacus type instrument or arrangement was intended to overcome this difficulty. A small child can learn to add and subtract by one, and multiplication is not much harder. Division is, by its nature, less simple. In the book a few simple numerical exercises are worked out. 634×523 is broken up into

$$\begin{matrix}\mathrm{C\ X\ I}\\ 6\ 3\ 4\end{matrix}\times\begin{matrix}\mathrm{C\ X\ I}\\ 5\ 2\ 3\end{matrix}=\begin{matrix}\mathrm{C\ X\ I}\\ 6\ 3\ 4\end{matrix}\times\left[\begin{matrix}\mathrm{C\ X\ I}\\ 3\end{matrix}+\begin{matrix}\mathrm{C\ X\ I}\\ 2\end{matrix}+\begin{matrix}\mathrm{C\ X\ I}\\ 5\end{matrix}\right]$$

The worst example treated is 734 divided by 23.

A University library should have a copy. Its interest cuts across departments. Individual teachers will find in the book very many interesting items.

A. J. KEMPNER, University of Colorado

The Concept of Number. A Chapter in the History of Mathematics with Applications of Interest to Teachers. By Christoph J. Scriba, with the assistance of M. E. Dormer Ellis. Bibliographisches Institut, Mannheim/Zürich, 1968, 216 pp., 6.90 D.M. (paper) (Telegraphic Review, August-September. 1969).

This is one of a number of recent, small, inexpensive, paper-back mathematics books written primarily for use in graduate classes in a Teacher Education Program. The authors or editors of these books, who have been careful to respect the needs and the maturity of prospective readers, have skillfully integrated results of current mathematical research with traditional subject matter and treatment. Thus the student has a taste of the spirit of discovery, an attribbute which does much to enliven courses in a Teacher Education Program.

The present book is devoted to a discussion of material which can be found in books on the History of Mathematics, Modern Algebra, Higher Arithmetic, Theory of Numbers, etc. As the author says, there seems to be no other book in the English language which has gathered together in a single volume his selection of subject matter. The student will find the book interesting and well written; he will find it a challenge, though not on too advanced a level. Since chapter headings are given in the telegraphic review cited above, we will refer here only to some of the important features of the author's work.

After discussing the generally more familiar topics of the development of number systems and the elementary operations with numbers, the author goes into a discussion in Chapter IV of algebra and number theory. In this chapter he discusses the theory of congruences, Euclid's algorithm, and Diophantine arithmetic. Chapter V, the final chapter, is devoted to a discussion of contributions to the number concept since the beginning of the nineteenth century. Here the author discusses works of Hamilton and Cayley, Liouville, Hermite, Peano, Dedekind, and Cantor, Hilbert and others. What a pity, however, that the names of these mathematicians have been omitted from an otherwise carefully prepared index. The book is copiously illustrated by tables and diagrams; it contains a wealth of exercises and problems, with a section on answers to selected problems; and an extensive bibliography. With the exception of the classic references, Menninger's *Zahlwort und Ziffer*, Tropfke's *Geschichte der Elementar Mathematik*, and one or two others, the bibliography of forty titles is confined to books in the English language. An exceedingly helpful innovation is the brief description which follows the title of an individual item listed in the bibliography.

Unfortunately the process of photographic reproduction and reduction in size, by means of which the book has been published, has rendered the print so small that reading the text is difficult. One hopes that a wide acceptance of this fine book will warrant a more worthy form of publication of a new edition in the near future.

LAURA GUGGENBUHL, Hunter College

TELEGRAPHIC REVIEWS

Telegraphic reviews are intended to give prompt notice of books, with sufficient information to assist in deciding whether to order an examination copy or suggest library purchase. Possible uses are indicated as follows:

B = college bookstore stock L = library purchase
P = professional reading S = supplementary reading
T = textbook E = teacher education
13 to 18 = freshman to second year graduate level usage
1 to 4 = approximate time in semesters to cover text
* = positive emphasis ? = negative emphasis

Books on high-school material (pre-calculus) are denoted REMEDIAL, and normally receive telegraphic reviews only if they are written for college students. Publishers are denoted by the standard abbreviations used in *Books in Print*, which gives complete addresses.

ALGEBRA-GROUP THEORY, T(18: 1), P, L. *Infinite Abelian Group Theory: Chicago Lectures in Mathematics*. Phillip A. Griffith. U of Chicago Pr, 1970, 152 pp, $2.50 (P). "The main purpose ... to arrive at P. Hill's version of Ulm's Theorem along with ... homological techniques in dealing with the notion of purity." Exercises. L.A.S.

ALGEBRA-NUMBER THEORY, P, L. *Number Theory: Colloquia Mathematica Societatis Janos Bolyai, 2*. Ed: P. Turán. North-Holland, 1970, 244 pp, $14. Twenty-one papers from a colloquium on number theory in Debrecen, Hungary, 1968, and a section of open questions posed by participants. L.A.S.

ALGEBRAIC NUMBER THEORY, T(18: 2), P. *Algebraic Number Theory*. Serge Lang. A-W, 1970, 354 pp, $14.95. A graduate text on algebraic and analytic number theory, written to supersede Lang's *Algebraic Numbers* and containing a good deal more material (e.g. class field theory) than that book. No problems. J.D.-B.

ANALYSIS, T(18: 2), P, L. *Classical and Modern Integration Theories*. Ivan N. Pesin (L'Vov U, U.S.S.R.). Transl: Samuel Kotz. Acad Pr, 1970, 195 pp, $12.50. This book "presents a detailed historical survey of the development of classical integration theory from Cauchy to Daniell..., (and) places the ideas of the classical authors (e.g. Borel, Lebesgue, Young, Stieltjes, and Radon) in the perspective of modern mathematics." The translation and editing make Pesin's work highly enjoyable reading. Contains an extensive bibliography. T.A.V.

ANALYSIS, MEASURE THEORY, E, S. *Notions de Mêsures et Nombres Réels*. Lucienne Félix (Agrégée de l'Université). Librairie Scientifique Albert Blanchard, Paris, 1970, 107 pp, $3,27. Written for teachers in French academic high schools, this book deals with the real numbers and (Borel) measure in R^n. It is intended not as a text but as an aid to deeper understanding of the calculus. It can perform the same function for college teachers of elementary calculus, though little in it will be entirely new to them. The book appears to be carefully written, though no attempt is made to prove everything. However, no satisfactory definition of the real numbers is given, and the Peano square--filling curve is

called a simple closed curve. J.D.-B.

ANALYTIC NUMBER THEORY, P, L. *Modular Functions in Analytic Number Theory.* Marvin I. Knopp (U of Ill). Markham, 1970, 150 pp, $11.50. A treatise on applications of the theory of modular forms to the study of the number of partitions of a positive integer into positive integers and the number of ways of representing a positive integer as a sum of squares. It is meant to be self-contained for a reader who has had a good first-year course in analysis, but that may be a bit optimistic. J.D.-B.

CHARACTERISTIC ROOTS AND VECTORS, T(15-16), S, L(UNDERGRAD). *Latent Roots and Latent Vectors.* S.J. Hammarlang. U of Toronto Pr, 1970, 172 pp, $13.50. This text, with exercises and solutions, treats the special topic given by its title. The subject is rapidly developed in a well-written exposition emphasizing methods for finding characteristic roots and vectors. There is a chapter treating applications of latent roots and vectors. R.J.

CYBERNETICS, P, L. *Introduction to Economic Cybernetics.* Oskar Lange. Pergamon Pr, 1970, 183 pp, $6. A translation from Polish. Deals with general cybernetic principles and mathematical formulations in economic planning in socialist countries. Valuable for those interested in contemporary cybernetics. Contains bibliography of 63 titles, about half in English. K.W.

DIFFERENTIAL EQUATIONS, P, L. *Decay of Solutions of Systems of Nonlinear Hyperbolic Conservation Laws.* James Glimm and Peter D. Lax. *Memoirs of the American Mathematical Society, Number 101.* AMS, 1970, 112 pp, $2.10 (P). A study of a strictly hyperbolic system of two nonlinear conservation equations, with results about decay of solutions for certain kinds of initial data. D.F.A.

DIFFERENTIAL EQUATIONS, T(15-16). *Topics in Ordinary Differential Equations: A Potpourri.* William D. Lakin and David A. Sanchez. Prindle, 1970, 154 pp, $5.95 (P). Assumes minimal background of first course in differential equations. The focus is on asymptotic expansions and perturbation methods, regular and singular. Material on the boundary value problem. T.A.V.

ECONOMICS, S, P, *L. *Selected Readings in Econometrics from Econometrica.* John W. Hooper and Marc Nerlove. M.I.T. Pr, 1970, 498 pp, $15. Twenty-two articles from 1932 to 1964 which have played a major role in the development of econometrics. The articles are divided equally between methodology and applications. Useful as a reference work. W.C.R.

EDUCATION, STATISTICS AND PROBABILITY, E, *P, L. *K-13 Mathematics: Some Non-Geometric Aspects. Part I: Statistics and Probability.* K-13 Arithmetic-Algebra Study Committee, Ontario Institute for Studies in Education, 1970, 56 pp, $2.50 (P). This report gives some ideas on curriculum development and objectives of mathematics instruction in general, and then makes some specific recommendations and detailed suggestions for integrating material on statistics and probability into the K-13 mathematics curriculum. Good bibliography. Part II will deal with computation and logic. R.S.K.

GENERAL, S. *Prepare Now for a Metric Future.* Frank Donovan. Weybright and Talley, 1970, 212 pp, $5.95. History and comparison of various world systems of weights and measures. Advocates U.S. conversion to metric system, giving pros and cons. Light and informative. Suitable for high school onward. A.G.

GENERAL, P, L. *Gesammelte Abhandlungen.* David Hilbert. Springer-Verlag, 1970, $27 (3 vols), consisting of *Zahlentheorie, Vol. I,* 539 pp; *Algebra-Invariantentheorie-Geometrie, Vol. II,* 453 pp; *Analysis-Grundlagen der Mathematik-Physik-Verschiedenes-Lebensgeschichte, Vol. III,* 435 pp. Though it does not say so, this is apparently a republication of the original edition, first published between 1932 and 1935. J.D.-B.

GEOMETRY, *T(15-16), E, *S, P, *L. *A Course of Geometry For Colleges and Universities.* D. Pedoe. Cambridge U Pr, 1970, 449 pp, $11.50. A heartening sign of a long-needed revival in the teaching of geometry. Includes vector proofs of Euclidean and affine theorems, theory of circles, mappings of the Euclidean plane, and projective geometry. Presumes some acquaintance with vector spaces, the solution of sets of homogenous equations, and matrix notation. J.N.C.

GEOMETRY FOR SCIENTISTS, P (SCIENTISTS, ENGINEERS), S, L. *Space Through The Ages.* Cornelius Lanczos. Acad Pr, 1970, 320 pp, $11.50. Traces the development of geometry from its central role in Greek education to the discoveries of Einstein and the formulation of Hilbert's Function Space, events which "brought geometry once more back to its ancient glory." In preparation for the chapters on Einstein's theory of gravitation and abstract spaces, it includes chapters on tensor algebra and analyses and the geometry of Gauss and Riemann. Presumes elementary calculus. J.N.C.

INTEGRAL EQUATIONS, P, L. *Volterra Integral Equations and Topological Dynamics.* Richard K. Miller and George R. Sell. *Memoirs of the American Mathematical Society, Number 102.* AMS, 1970, 67 pp, $1.80 (P). Discusses the basic theory of local dynamical systems, and uses this to study bounded solutions of Volterra equations and to characterize their ω-limit sets. Gives a number of applications, including ones to ordinary differential equations and asymptotic behavior of solutions of Volterra equations. D.F.A.

LINEAR ALGEBRA, T(13-14: 1). *Linear Algebra.* D.C. Murdoch. Wiley, 1970, 312 pp, $9.95. Good basic text on Linear Algebra. Revised from a 1957 edition with a more abstract approach to fit the needs of today's students. Still an elementary course designed for students who need a knowledge of linear algebra, matrices and their applications. Includes determinants, linear transformations, diagonalization theorems, reduction of quadratic forms, inner product spaces. L.L.K.

LINEAR PROGRAMMING, S, P, L. *User's Guide to Linear Programming.* Hans G. Daellenbach and Earl J. Bell. P-H, 1970, 226 pp, $7.95. Problem formulation and computer solutions. R.W.N.

LOGIC, S, P, L. *Intuitionism and Proof Theory.* A. Kino, J. Myhill and R.E. Vesley. North-Holland, 1970, 516 pp, $28. This volume constitutes the proceedings of the Conference on Intuitionism and

Proof Theory held at the State University of New York at Buffalo in August 1968. It contains thirty-one papers by distinguished contributors to these increasingly important areas of foundational studies. Some of these papers are accessible to nonspecialists and effectively point to new directions currently being taken in these branches of mathematics. L.C.L.

MATHEMATICAL ECONOMICS, T(14-17), S, L. *Introduction to Sets and Mappings in Modern Economics*. Hukukane Nikaido. Transl: Kazuo Sato. North-Holland, 1970, 343 pp, $20.75. Presupposes calculus but not any economics. Linear algebra, convex sets, fixed point theorems, and applications to linear models, optimization problems, saddle point problems, and equilibrium in economics. Translation from the Japanese. Note the uneconomical price. F.L.W.

NUMBER THEORY, P, L. *Zahlentheorie*. Paul Bachmann. Johnson Repr, 1968, $85 (6 vols). An unchanged reprinting of Bachmann's six-volume treatise on number theory, consisting of *Die Elemente der Zahlentheorie, Vol. I* (1892), 264 pp; *Die Analytische Zahlentheorie, Vol. II* (1894), 494 pp; *Die Lehre von der Kreistheilung, Vol. III* (1872), 299 pp; *Die Arithmetik der Quadratischen Formen, Vol. IV, Part 1* (1898), 668 pp; *Die Arithmetik der Quadratischen Formen, Vol. IV, Part 2* (1923), 537 pp; *Allgemeine Arithmetik der Zahlenkörper, Vol. V* (1905), 548 pp. J.D.-B.

NUMBER THEORY, T(15-17), L. *An Introduction to the Theory of Numbers*. Ralph G. Archibald. Merrill, 1970, 305 pp, $8.95. More inclusive than the usual text in this field. Also provides more than the usual number of exercises and problems. A series of useful notes on each chapter. K.W.

OPTIMIZATION, *T(15-17: 1, 2), P, L. *Introduction to Methods of Optimization (NB)*. Leon Cooper and David Steinberg. Saunders, 1970, 381 pp, $12.50. Introductory chapters on matrix algebra and n-dimensional geometry. Chapters on classical and discrete techniques, linear, nonlinear, and integer programming. Also many examples and exercises. R.W.N.

OPTIMIZATION, CONVEXITY, *P, L. *Convexity and Optimization in Finite Dimensions I. Grundlehren der Math. Wissenschaften, Volume 163*. Josef Stoer and Christoph Witzgall. Springer-Verlag, 1970, 293 pp and 19 figures, $14.90. Monograph on linear inequalities, polyhedra, convex sets and convex functions providing the theoretical background for a planned second volume on algorithms for optimization problems. Bibliography of 17 pages. R.W.N.

ORDINARY DIFFERENTIAL EQUATIONS, NUMERICAL ANALYSIS, T(15-17: 1), S, *P, L. *Computation and Theory in Ordinary Differential Equations*. James W. Daniel and Ramon E. Moore. Freeman, 1970, 172 pp, $7.50. Assumes some ordinary differential equations and numerical analysis. Analyzes the computational efficiency of methods for initial and boundary value problems. 60 pages on the use of coordinate transformations. R.W.N.

PARTIAL DIFFERENTIAL EQUATIONS, T(18: 2), *P, L. *Linear Differential Operators with Constant Coefficients*. V.P. Palamodov. Transl: A.A. Brown. Springer-Verlag, 1970, 443 pp, $27. The book consists of two parts. The first discusses cohomological and

algebraic structures of spaces of analytic functions of several variables. The second contains a systematic development of the theory of systems of partial differential equations with constant coefficients. The translation is concise and very readable. Contains an extensive bibliography. T.A.V.

PHILOSOPHY OF MATHEMATICS, S, *L, *Essays on Bertrand Russell.* E.D. Klemke. U of Ill Pr, 1970, 458 pp, $10.95. 26 Essays on Russell's ontology, theories of reference, and philosophy of mathematics. 10 of them appear here for the first time. F.L.W.

PROBLEMS, T(13-15), S, L, *Aufgabensammlung Zur Höheren Mathematik, Band I.* N.M. Günter and R.O. Kusmin, VEB Deutscher Verlad, der Wissenschaften, Berlin, 1968, 508 pp. First of two volumes of the sixth edition of a translation of a Russian book of problems. Over 4300 problems, ranging from routine to difficult, on plane and solid analytic geometry, calculus, algebra and differential equations, with statements of some definitions and theorems and solutions of most problems. J.D.-B.

REMEDIAL, ALGEBRA AND ANALYTIC GEOMETRY, T+++(2), *Intermediate Algebra.* Edward Gaughan. Brooks/Cole, 1970, 438 pp, $8.50. Standard topics in high school algebra, the elements of functions and analytic geometry, plus a chapter on probability. Well written, with many exercises. A.G.

RIEMANNIAN GEOMETRY, T(18: 1), P, *L, *Integral Formulas in Riemannian Geometry.* Kentaro Yano. Marcel Dekker, 1970, 156 pp, $10.75. Some of the important global results of Riemannian geometry are obtained by applications of integral formulas such as those used in classical differential geometry. Extensive problems and bibliography. W.C.R.

STATISTICS, S?, *Statistical Functions.* Buddy L. Myers and Norbert L. Enrick. Kent S U Pr, 1970, 174 pp, $8, $4.25 (P). Intended to provide detailed derivations of the principal statistical functions and formulas, assuming only a background of elementary algebra and basic calculus. Typographical errors, mathematical inaccuracies and other inconsistencies make this book of doubtful value. R.S.K.

STATISTICS, T(14: 2), *Statistics: Probability, Inference, and Decision.* William L. Hays and Robert L. Winkler. Two volumes. HR & W, 650 pp, $11.50; 320 pp, $9.50. Volume I covers basic probability theory (from a set-theoretic point of view) and statistical inference and decision, with extensive coverage of Bayesian inference and decision theory. Volume II contains material on regression and correlation, analysis of variance, and non-parametric methods. The level of the book is above that of elementary texts. Some calculus is used and theory is emphasized. New concepts are introduced both mathematically and verbally and many proofs are given for discrete random variables. The problem sets are good and fairly extensive tables are included. R.S.K.

STATISTICS, T(16-17: 1), S, P, L, *Order Statistics.* H.A. David. Wiley, 1970, 272 pp, $13.95. Distribution theory, estimation and hypothesis testing, the treatment of outliers, asymptotic theory. Extensive bibliography. F.L.W.

STATISTICS, E, *S, *B, *L. *Programmed Statistics, with Chapters on Probability, Computer Theory, and Programmed Instruction.* Richard Bellman, John C. Hogan, and Ernest M. Scheuer. HR & W, 1970, 115 pp $3.10 (P). Designed to give that basic information about probability and statistics that every high school and elementary school teacher should know, and uses no mathematics beyond arithmetic and elementary algebra. Material on statistics is programmed, but conventionally written chapters discuss elementary probability theory, digital computers, programmed learning, tables of random digits, and a new computational aids for educators. Contains many references and a glossary of computer and programmed instruction terms. The book is ideally suited for self-study and as a supplement to courses in education and educational psychology.D.F.A.

STATISTICS, TIME SERIES, T(17-18: 1, 2), P, L. *Time Series Analysis: Forecasting and Control.* George E.P. Box and Gwilym M. Jenkins. Holden-Day, 1970, 553 pp, $24. Concerned with the building (as opposed to just the fitting) of statistical models for discrete time series, and the use of these models in forecasting and control. It contains many illustrative figures and tables and includes a section describing useful computer programs. No exercises. R.S.K.

STATISTICS, TIME SERIES, T(18), P, L. *Multiple Time Series.* E.J. Hannan. Wiley, 1970, 536 pp, $21.95. Book in the Wiley Series in Probability and Mathematical Statistics. Concerned with the theory of statistical analysis of time series data -- essentially no numerical examples. In particular, develops the theory for the case in which multiple measurements are made at each time point. Good bibliography of recent material. R.S.K.

SYSTEMS ENGINEERING, T(16: 1), S, P, L. *Optimization and Probability in Systems Engineering.* John G. Rau. Van Nostrand, 1970, 403 pp, $15. Math modeling. Optimization and probabilistic techniques with engineering applications. R.W.N.

Reviewers Whose Initials Appear Above

David F. Appleyard, Carleton; John Dyer-Bennet, Carleton; Judith N. Cederberg, St. Olaf; Arthur Gropen, Carleton; Richard Jarvinen, Carleton; Lorraine L. Keller, St. Olaf; Richard S. Kleber, St. Olaf; Loren C. Larson, St. Olaf; R.W. Nau, Carleton; William C. Ramaley, Carleton; Linda A. Seebach, St. Olaf; T.A. Vessey, St. Olaf; Kenneth Wegner, Carleton; Frank L. Wolf, Carleton.

NOTABLE

Microfilm copies of a table of prime factors for numbers to 10^8 are available from Reproduction Systems, Inc., 1399 S. 700 E., Salt Lake City, Utah 84105, for about $50. These were prepared by Kay Litchfield, an undergraduate at Brigham Young University.

The Journal of the Royal Statistical Society, Volumes 1-89, 1838-1926, is being reprinted by Wm. Dawson and Sons Ltd., Cannon House, Folkestone, Kent, England. This work is of as much interest to social scientists as to mathematicians since it contains huge quantities of raw data.

NEWS AND NOTICES

EDITED BY RAOUL HAILPERN, SUNY at Buffalo

Readers are invited to contribute to the general interest of this Department by sending news items to Mathematical Association of America, 1225 Connecticut Avenue, N.W., Washington, D.C. 20036. Items must be submitted at least two months before publication can take place.

PERSONAL ITEMS

University of Missouri—Columbia: Mr. K. R. Pierce, University of Wisconsin, Rock County Center, has been appointed Assistant Professor; Professor R. R. Summerhill is on leave at the Institute for Advanced Study, Princeton.

Associate Professor H. J. Arnold, Oakland University, has been promoted to Professor and appointed Acting Chairman of the Department of Mathematics.

Professor B. H. Bissinger, Lebanon Valley College, has been appointed Coordinator of Mathematical Sciences at the Capitol Campus of the Pennsylvania State University.

Dr. Eleazer Bromberg, New York University, has been named Vice Chancellor for Academic Affairs.

Mr. W. E. Brown, Dartmouth College, has been appointed Assistant Professor at the University of the Pacific.

Dr. Benjamin Epstein has been appointed Professor at the Technion—Israel Institute of Technology, Haifa, Israel.

Assistant Professor R. A. Groeneveld, Mount Holyoke College, has been appointed Associate Professor at Iowa State University.

Dr. M. C. Hartley, Chairman of the Department of Mathematics at the University of Tampa, was presented the G. Truman Hunter Award as the outstanding faculty member for the 1969–70 academic year.

Professor I. N. Herstein of the University of Chicago has accepted a position with the Weizmann Institute in Israel. He will hold a joint appointment at the Weizmann Institute and at the University of Chicago, spending the winter and spring quarter in Israel and the rest of the time in Chicago.

Dr. A. E. Hoffman, State University College at Cortland, has joined the Mathematics Staff of the State University College at Geneseo.

Professor Emeritus C. H. Yeaton, Oberlin College, died on July 7, 1970 at the age of 84. He was a Charter Member of the Association.

ANNUAL REPORT OF THE NATIONAL RESEARCH COUNCIL

The Division of Mathematical Sciences, National Research Council, calls attention to its Annual Report for 1969–1970. The Report contains an address by Professor I. N. Herstein, University of Chicago, and statements by six graduate students which made up a Symposium on Graduate Education in Mathematics held on the evening of March 16.

You may have a copy of the report by writing to the Division of Mathematical Sciences, National Research Council, 2101 Constitution Avenue, N. W., Washington, D. C. 20418.

R. L. MOORE INSTRUCTIONAL TECHNIQUES

John G. Harvey and Douglas R. Forbes of the Department of Mathematics, University of Wisconsin, are making an in-depth study of the R. L. Moore instructional techniques, sometimes called "the Texas Method." They would appreciate receiving information from any mathematicians who have tried similar methods of instruction. Please write to PROFESSOR JOHN G. HARVEY, DEPARTMENT OF MATHEMATICS, 213 VAN VLECK HALL, UNIVERSITY OF WISCONSIN, MADISON, WISCONSIN 53706.

MU ALPHA THETA MATHEMATICAL BOOKLIST

The National High School and Junior College Mathematics Club, Mu Alpha Theta, has revised and updated its list of *Enrichment Mathematics Books for School and Public Libraries*. Single copies of the 1970 list may be obtained without charge by sending a self-addressed stamped number 10 (business size) envelope to: High School Book List, Mu Alpha Theta, Mathematics Department, University of Oklahoma, 1000 Asp Avenue, #215, Norman, Oklahoma 73069. Persons desiring quantities of the list are encouraged to send $10 per hundred copies to help with the cost of the project.

CBMS NEWSLETTER SUBSCRIPTIONS NOW AVAILABLE

The Conference Board of the Mathematical Sciences announces that subscriptions to its *Newsletter*, beginning with the January 1971 issue, are available at a special rate of $2.00 per year to individuals who belong to one or more of the member-societies of the Conference Board—namely, the American Mathematical Society, the Association for Computing Machinery, the Association for Symbolic Logic, the Institute of Mathematical Statistics, the Mathematical Association of America, the National Council of Teachers of Mathematics, the Operations Research Society of America, the Society for Industrial and Applied Mathematics, the Society of Actuaries, and The Institute of Management Sciences. In authorizing this action, the Conference Board's Council, representing the above professional societies, has also authorized the Conference Board to make subscriptions to its *Newsletter* available to other individuals, and to libraries and institutions, at a regular rate of $4.00 per year.

The *Newsletter* of the Conference Board of the Mathematical Sciences seeks to present up-to-date news and information of direct interest and relevance to the broad mathematical community represented by its member-societies, including such items as current programs and plans of agencies of the Federal Government and private foundations affecting the mathematical sciences; reports on recent or coming national or international mathematical events of note; honors and awards to individual mathematical scientists and new appointments to important posts; activities and announcements of broad interest on the part of the various mathematical professional organizations.

Beginning in 1971, the CBMS *Newsletter* is to be published in four sixteen-page issues per year, in January, March, May and October. At least for calendar 1971 the *Newsletter* will continue to be distributed free of charge to department chairmen in the mathematical sciences in universities and four-year colleges and to heads of mathematical research groups in industry, as well as to officials of CBMS member-organizations and of various governmental agencies. Orders for subscriptions to the *Newsletter* may be addressed to CBMS, 2100 Pennsylvania Ave., N.W., #834, Washington, D.C. 20037. Individual orders at the special $2.00 rate should be accompanied by pre-payment and should list those member-organizations of CBMS to which the individual belongs.

MATHEMATICAL ASSOCIATION OF AMERICA

Official Reports and Communications

NEW GROUP LIFE INSURANCE PLAN

A Group Life Insurance Plan has been instituted for our members in early January. This Plan will provide $10,000 group term life insurance coverage, *regardless of health*, for members under age 50 who have been regularly at work at least 30 hours a week for 30 days immediately prior to enrollment. Members age 50 through 69 may also request

coverage, subject to insurance company acceptance on the basis of health questions. Coverage for dependents in lesser amounts will also be available.

Members may request additional amounts of coverage of $10,000, $20,000, or $30,000, subject to acceptance by the insurance company. All coverage reduces by 50% at age 65 and terminates after age 70, when a permanent individual insurance policy can be obtained without regard to health.

Members in Ohio, Texas and Wisconsin will be eligible for individual policies with comparable coverage, in accordance with the insurance laws of these states. (Texas law, however, requires that its residents submit a health statement to the insurance company for review.)

The Plan will also be available to members of AMS and SIAM who will be covered under the same group policy. A member is therefore limited to $10,000 coverage regardless of health and a maximum of $30,000 additional coverage even if he or she also belongs to AMS or SIAM or both.

Continental Assurance Company, of Chicago, Illinois will underwrite the new Plan. Established in 1911, this Company now ranks among the top 2% of U.S. insurance companies. If there is adequate participation and normal mortality for the Plan, experience credits will be used to reduce premium or increase benefits, or both. Credits will be used solely for the benefit of insured members and no part of any credit will be paid to or used for the benefit of the Association. These credits cannot, of course, be guaranteed.

The key personnel of the Administrator of the Plan have successfully managed a Group Insurance Program for the American Society of Civil Engineers for the past 20 years. They also administer Insurance Programs for the American Statistical Association, the Association for Computing Machinery, The Institute of Electrical and Electronics Engineers, Inc., and many other scientific and technical associations. They assure us that at no time will our members be approached through personal solicitation.

The Plan will be offered to members at rates substantially lower than the cost of similar coverage on an individual basis. The success of the Plan will depend upon the interest and active support of our membership. While MAA has approved this Plan for its members, no expense will be incurred by the Association in developing or continuing it. We believe that many members will consider this a worthwhile project, and the Plan is being offered for their careful consideration.

MAA SUMMER SEMINAR

The Mathematical Association of America will conduct a Summer Seminar in the theory of probability and mathematical statistics to be held at Williams College from 21 June to 30 July, 1971. The staff for the seminar will be Professor Frank Spitzer, Cornell University, and Professor Geoffrey Watson, Princeton University. The program, funded by the National Science Foundation, is being conducted for teachers of probability or statistics in colleges and universities which offer an undergraduate major in mathematics but which do not have a Ph.D. program. Participants will be expected to hold the Ph.D. degree or have comparable qualifications. Further information can be obtained by writing Prof. Neil R. Grabois, Director of the 1971 Summer Seminar, Williams College, Williamstown, Mass. 01267.

Correction (*April Meeting of the Metropolitan New York Section*). The report of the April meeting of the Section on page 922 of the October 1970 issue of this MONTHLY implies that the Sectional Governor is Professor Abraham Schwartz. This is an error; the Governor from the Metropolitan New York Section is Professor Gerald Freilich of City College of New York. At the meeting, the report of the Sectional Governor was read by Professor Schwartz in Professor Freilich's absence.

CALENDAR OF FUTURE MEETINGS

Fifty-second Summer Meeting, Pennsylvania State University, University Park, August 30–September 1, 1971.

Fifty-fifth Annual Meeting, Las Vegas, Nevada, January 19–21, 1972.

The following is a list of the Sections of the Association with dates of future meetings so far as they have been reported to the Editorial Director.

ALLEGHENY MOUNTAIN, Geneva College, Beaver Falls, Pennsylvania, May 7–8, 1971.

FLORIDA, Florida Southern College, Lakeland, March 19–20, 1971.

ILLINOIS, Eastern Illinois University, Charleston, May 14–15, 1971.

INDIANA

IOWA, Loras College, Dubuque, April 23, 1971.

KANSAS

KENTUCKY, Western Kentucky University, Bowling Green, April 2–3, 1971.

LOUISIANA-MISSISSIPPI

MARYLAND-DISTRICT OF COLUMBIA-VIRGINIA

METROPOLITAN NEW YORK, Nassau Community College, Long Island, April 3, 1971.

MICHIGAN, Western Michigan University, Kalamazoo, May 7–8, 1971.

MISSOURI, Missouri Southern College, Joplin, April 30-May 1, 1971.

NEBRASKA, Nebraska Wesleyan University, Lincoln, April 30-May 1, 1971.

NEW JERSEY

NORTH CENTRAL, University of Minnesota, Minneapolis, May 8, 1971.

NORTHEASTERN

NORTHERN CALIFORNIA

OHIO

OKLAHOMA-ARKANSAS, University of Tulsa, Tulsa, March 12–13, 1971.

PACIFIC NORTHWEST, Oregon State University, Corvallis, June 18–19, 1971.

PHILADELPHIA

ROCKY MOUNTAIN, Weber State College, Ogden, Utah, May 7–8, 1971.

SOUTHEASTERN, University of Alabama, Tuscaloosa, March 26–27, 1971.

SOUTHERN CALIFORNIA, San Fernando Valley State College, Northridge, March 13, 1971.

SOUTHWESTERN, Arizona State University, Tempe, April 2–3, 1971.

TEXAS, Midwestern University, Wichita Falls, April 16–17, 1971.

UPPER NEW YORK STATE

WISCONSIN, Ripon College, Ripon, April 30-May 1, 1971.

FUTURE MEETINGS OF OTHER ORGANIZATIONS

AMERICAN ASSOCIATION FOR THE ADVANCEMENT OF SCIENCE, Philadelphia, December 26–31, 1971.

AMERICAN MATHEMATICAL SOCIETY, Pennsylvania State University, University Park, August 31-September 3, 1971.

AMERICAN SOCIETY FOR ENGINEERING EDUCATION, U. S. Naval Academy, Annapolis, June 21-24, 1971.

ASSOCIATION FOR COMPUTING MACHINERY, Los Angeles, California, March 23, 1971.

ASSOCIATION FOR SYMBOLIC LOGIC, Beverly Hilton Hotel, Los Angeles, March 25–26, 1971.

CENTRAL ASSOCIATION OF SCIENCE AND MATHEMATICS TEACHERS, Detroit, Michigan, November 25–27, 1971.

FIBONACCI ASSOCIATION, University of San Francisco, April 24, 1971.

INSTITUTE OF MATHEMATICAL STATISTICS

MU ALPHA THETA

NATIONAL COUNCIL OF TEACHERS OF MATHEMATICS, Anaheim, California, April 14–17, 1971.

OPERATIONS RESEARCH SOCIETY OF AMERICA, Sheraton Dallas, Dallas, May 5–7, 1971.

PI MU EPSILON, Pennsylvania State University, University Park, August 31-September 1, 1971.

SOCIETY FOR INDUSTRIAL AND APPLIED MATHEMATICS, Seattle, Washington, June 28–30, 1971.

1971 Macmillan

PRINCIPLES OF ARITHMETIC AND GEOMETRY FOR ELEMENTARY SCHOOL TEACHERS

By CARL B. ALLENDOERFER, University of Washington

This text presents a complete discussion of the structure of the number system and a survey of informal geometry. Each topic is considered at the intuitive, theoretical, and practical levels. Programmed exercises, a Readiness Test preceding each chapter, a Post Test following each chapter, and a Summary Test at the end of each section are particularly helpful. Set theory is used extensively throughout. Thirteen brief sound and color films, which correspond to approximately one-half of the text, are available.

Instructor's Manual, *gratis.*

1971 approx. 608 pages $9.95

INTERMEDIATE ALGEBRA

By WARD D. BOUWSMA, Southern Illinois University

Intended for algebra courses in both junior colleges and four year schools, this text covers the basic concepts of algebra with unusual thoroughness. Discovery-oriented discussions provide the student with the ability to devise creative approaches to problem solving.

Combined Instructor's and Solutions Manual, *gratis.*

1971 approx. 320 pages $8.50

ESSENTIALS OF TRIGONOMETRY

By IRVING DROOYAN, WALTER HADEL, and CHARLES C. CARICO, all of Los Angeles Pierce College

The topics in this book for a one-semester course in trigonometry are arranged to meet the needs of students also enrolled in courses that make early use of trigonometric ratios and their applications.

Teacher's Manual, *gratis.*

1971 approx. 320 pages $8.95

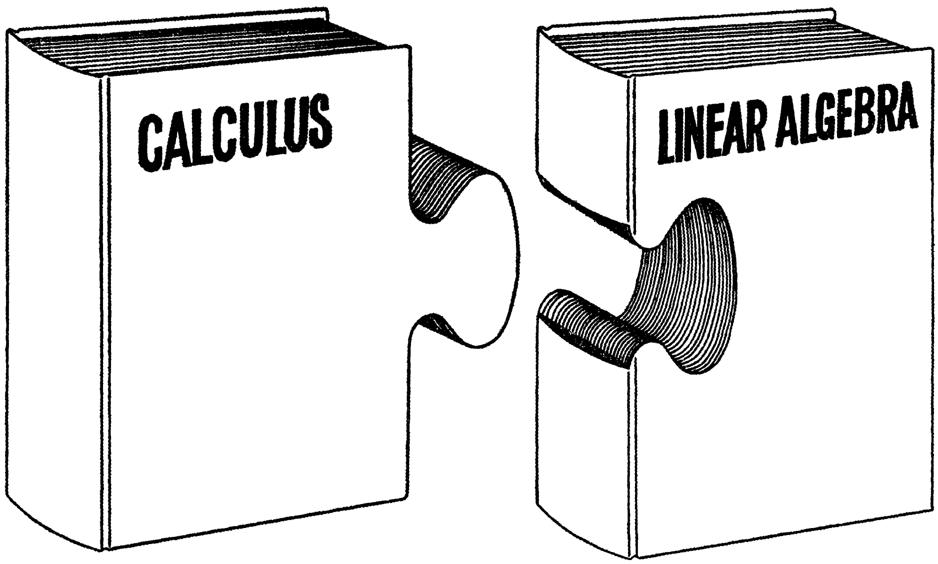
BASIC MATHEMATICS WITH ELECTRONIC APPLICATIONS

By JULIUS L. SMITH, Collins Radio Company,
and DAVID S. BURTON, Chabot College

This is the only completely up-to-date book that successfully integrates mathematical principles and electronic applications, while requiring a minimal background in high school arithmetic.

Instructor's Manual, *gratis.*

1971 approx. 640 pages prob. $11.95


CALCULUS
LINEAR ALGEBRA

GEORGE BANTA COMPANY, INC., MENASHA, WISCONSIN

THE AMERICAN MATHEMATICAL MONTHLY

(Founded in 1894 by Benjamin F. Finkel)

THE OFFICIAL JOURNAL OF

THE MATHEMATICAL ASSOCIATION OF AMERICA, INC.

VOLUME 78 NUMBER 3

CONTENTS

MARCH 1971

NOTICE TO AUTHORS

Specialized research is usually unsuitable; see *Statement of Policy* (vol. 76, p. 2). Manuscript preparation: Please use the *Manual for Monthly Authors* (vol. 78, p. 1) and follow the format in current issues of the MONTHLY. Manuscripts should be typewritten, triple-spaced with wide margins; submit two copies and keep one for protection against loss. Each author receives 50 free reprints.

Backlog: Main Articles 9 months, Math. Notes 9 months, Research Problems 7 months, Classroom Notes 8 months, Math. Education 6 months.

EDITORIAL CORRESPONDENCE AND MAIN ARTICLES: to HARLEY FLANDERS, American Mathematical Monthly, Tel Aviv University, Ramat Aviv, Israel (see Notice, vol. 77, 1970, p. 555); NOTES, etc.: to the corresponding Associate Editor; ADVERTISING CORRESPONDENCE: to RAOUL HAILPERN, Mathematical Association of America, SUNY at Buffalo, Buffalo, N. Y. 14214; CHANGE OF ADDRESS and SUBSCRIPTIONS: to A. B. WILLCOX, Mathematical Association of America, 1225 Connecticut Ave., N.W., Washington, D. C. 20036.

Annual dues for members of the Association (including a subscription to the American Mathematical Monthly) are $12.50. For nonmembers the subscription price is $18.00.

PUBLISHED BY THE ASSOCIATION at Washington, D. C., and Menasha, Wisconsin during the months of January, February, March, April, May, June–July, August–September, October, November, December.

Second-class postage paid at Washington, D. C., and additional mailing offices. Acceptance for mailing at special rate of postage provided for in the Act of February 28, 1925, embodied in Paragraph 4, Section 538, P. L. and R., authorized April 1, 1926.

PRINTED IN THE UNITED STATES OF AMERICA

OLD CAMBRIDGE DAYS

LEONARD ROTH†

The town of Cambridge is a rather insignificant little place situated some fifty miles to the north of London. Its only title to distinction is, and always has been, that it happens to be the site of a university. Why this should be so is a complete mystery; the historians have never been able to throw any light on it and probably never will. But how the university actually came about—that is known with some precision. The earliest universities of Christian Europe were all born in the same way: a great scholar settled in a town and attracted some students to his lectures; these disciples then stayed with him and, in their turn, began teaching the young men who continued to arrive; and so, without any forethought, the thing was done. Thus the University of Bologna, the first of the great European foundations, came into being simply because some of the local monks began giving public lectures on Roman law: this was about 1150. Today, in one of the principal squares of Bologna, there is a monument to those men who were in effect the first university lecturers of medieval Europe. (Throughout the centuries, many of Italy's greatest advocates have received their training in the Bologna law school.) The universities of Paris and of Oxford were both founded about fifty years later, in exactly the same way.

Now the universities of the second wave of foundations arose in a quite different manner. When learned men have been in one another's company for a sufficient length of time they usually begin to quarrel. From the highest motives, naturally: doctrinal questions, difficult philosophical points, and the like. In those early universities, the disputes sometimes grew so acute that the entire body of scholars split into two factions; and then the dissenting party would go away and found a rival institution elsewhere, just as a swarm of bees will leave the parent hive. Several French and Italian universities were founded in this

† Leonard Roth was Andrew Mellon Professor of Mathematics at the University of Pittsburgh when he and his Italian wife, Marcella Baldesi, lost their lives on November 28, 1968 in a collision.

L. Roth, born in London in 1904, was educated at Clare College, Cambridge, England. He graduated in 1926, a "Wrangler" in the Mathematical Tripos, and was appointed at the Imperial College of Science and Technology at London University. With few interruptions, he worked there until 1967, first as Lecturer and then as Reader in Pure Mathematics.

A pupil of H. F. Baker, he was awarded a Rockefeller Research Fellowship at Rome University in 1929, where he established lasting and fruitful relationships with Italian geometers. In fact, his original work is almost entirely devoted to algebraic geometry following the methods of the Italian school. His main contributions concern birationality and unirationality questions on algebraic manifolds, as well as Abelian, pseudo-Abelian, and group varieties, and are of lasting significance. A complete list of his publications, more than 90 in number, which will be found at the end of an obituary by E. G. Togliatti appeared in *Boll. Un. Mat. Ital.*, (4) 3, 1970, pp. 326–332.

Leonard Roth was a very good lecturer and had a deep and widespread knowledge of the arts, music, and literature. His many-sided gifts and his charm are only partially apparent from the following paper found among other MSS left by him. It was probably not written for publication. Those who had the privilege of knowing him will always recall his more intimate endowments: his unusual kindness, unpretentiousness, and deeply-felt humanity. B. SEGRE.

way, by swarms from Paris and Bologna respectively. And Cambridge was founded, about the year 1200, by a swarm from Oxford.

Incidentally, this last migration had a curious aftermath. Until fairly recently, anyone who proceeded to a master's degree at Oxford had to sign a document affirming that he would never in any circumstances lecture to the little town of Stamford (Stamford, by the way, is on the road from London to York). The reason for this procedure was that some time later, a second migration left Oxford and set up a new center of learning at Stamford. However, the project came to nothing, and the dissentients soon returned home to Oxford. But the university authorities were so alarmed by these two flights that they took steps to prevent any more.

And so, by the end of the thirteenth century, we find two university institutions thriving in England, and situated—by Continental standards at any rate—almost within stone's throw of each other. However, although so near on the map, the climatic conditions which they enjoyed—or endured, as the case may be—were very different indeed. Oxford has a mild winter and is rather enervating on the whole. But Cambridge is much worse off; and in the early Middle Ages, when it lay on the edge of a huge undrained fen, the weather there must have been truly horrible. Even today, with its cold and damp winter, aggravated by a wind blowing across the European plain, it is nothing to joke about.

Now very soon after Oxford and Cambridge were founded they were confronted with serious problems of discipline. The presence of a large body of young men, in every respect alien to the local townsfolk, led to frequent disorders, the so-called "town and gown" riots, which often culminated in a murder. The situation was particularly galling to the magistrates because of the immunity frequently enjoyed by priests and clerks. This still persists in some institutions: thus, in Bologna university, if one student commits an offense against another which is normally punishable by law, the police are powerless to intervene. The authorities of both English universities sought a solution to the problem by putting their students into halls of residence, which later became known as colleges. The Head of the House (as he is still officially designated at Oxford) or Master of the College, to give him his usual modern title, was endowed with considerable powers over the inmates. His second in command was known as the Father of the College; he was *in loco parentis*—and he is the ancestor of the modern Dean. We shall soon see what his role was in the Cambridge scholastic world.

We may remark in passing that scarcely any other university saw fit to follow the example set by Oxford and Cambridge; this is exceedingly odd since one and all were beset by the same disciplinary problem. At both the English universities the college system gradually brought about fundamental changes in the administration. As the centuries passed, the central authority lost more and more to the individual colleges, until the latter became very nearly autonomous bodies, each of them responsible for the supervision and teaching of their own

students. This shift in power was to have significant consequences for the study of mathematics in Cambridge, which is our main theme here.

There was, however, one important right which the University still retained: that of holding the degree examination. This took the form of a disputation of strict syllogistic character, in which no other kind of reasoning was permitted; it was a three-cornered affair—examiner versus candidate, with the Dean acting as buffer between the contestants. Whenever his "son" was held up for a syllogism, the Dean would slip one in; if the candidate was unable to parry the argument, the Dean would come to his assistance. One beautiful feature of these medieval examinations was that no candidate ever failed; he might do very well, in which case he received the distinction *summa cum laude*, or his performance might be execrable. But in any case, he was awarded his degree; that is what Deans were for in those days.

Echoes of this ancient ceremony are still to be heard in the Cambridge examination system of today. Because the parties to the dispute used to sit on three-legged stools, the Cambridge examination is called a Tripos; moreover, in memory of the disputation, everyone who gains a first class in the Mathematical Tripos is designated a Wrangler.

In those remote times there was no hint that mathematics was destined to play the predominant part in Cambridge university studies which it was later to be assigned. The syllabus followed the lines of the normal medieval curriculum; geometry (but not mathematics), logic, philosophy, music, with specialist courses for students of law, divinity, or medicine. We have to wait until the mid-seventeenth century for the emergence of mathematics as a major discipline, that is to say, for the period in which Isaac Barrow occupied the mathematical chair.

How important is Barrow in the history of seventeenth-century science? The answer given to this question depends largely on one's nationality. The fact is, most of the European nations have their own horse in the Calculus Stakes: thus the French have Fermat and the Italians have Torricelli; if the Russians have not yet entered a candidate, it is only a question of time before they do. Any Englishman would, I fancy, put his money on Barrow; and there are sound, not patriotic, reasons for such a choice. It seems, on reading Barrow's lectures, that he has come nearer to the general notions of derivative and integral than have any of his predecessors. Now Newton was Barrow's pupil, and he absorbed Barrow's ideas.

Newton is, of course, the greatest of all Cambridge professors; he also happens to be the greatest disaster that ever befell not merely Cambridge mathematics in particular but British mathematical science as a whole. This is a fact which the historians do not care to dwell upon; if they mention it at all, they rarely give it due emphasis.

When Newton succeeded Barrow in the Cambridge chair, he had already lost interest in mathematical studies and had turned his intellect into theological and speculative channels. Moreover, he became caught up in public affairs as

well. When William and Mary came to the throne in 1688, the British coinage was in a terribly debased condition, and it was essential to commerce that this should be remedied as quickly as possible. For some years the government shrank before the magnitude and peril of the task; in the end, however, the recoinage was decided upon: Newton was called to London as Master of the Mint, and under his supervision the operation was performed with complete success.

These aspects of Newton's career—his indifference to mathematics and his absenteeism—may be termed his negative contributions to the ruin of British science. His positive contribution was far more serious: I refer to his quarrel with Leibniz concerning the origins of the infinitesimal calculus. Throughout his life, Newton shrank from every kind of controversy: by a strange irony his very reticence landed him in the bitterest dispute in mathematical history. The main facts are so well known that I need only allude to them briefly. For some time, there had been subterranean rumblings in the hitherto cordial relations between Newton and Leibniz. In 1715, however, matters came to a head; there then appeared a lengthy historical review of the work done in this field by the two rivals. The article was anonymous, but it is now generally believed to have been written entirely by Newton himself. After surveying the whole situation apparently with complete objectivity, it ends by summing up carefully but decisively against Leibniz, virtually accusing him of having stolen Newton's ideas. At this the Continental mathematicians immediately leapt to Leibniz's defense; at the same time the British mathematicians ranged themselves behind Newton. And there, for over a century, they stayed; their patriotism and solidarity were manifested in a boycott of European mathematics—and, most regrettably, during the very period when the modern science was in its full flood of development. So the Great Sulk went on; and the work of the Bernoullis, of Euler, Lagrange, Laplace, Gauss, and Cauchy remained for Britons a dead letter.

But this is not the whole story by any means; even when the period of official isolation (so to speak) was over, its disastrous consequences continued to make themselves felt. Although during the nineteenth century British applied mathematics made spectacular strides, pure mathematics was more or less neglected and, with the exception of Arthur Cayley, Great Britain produced no pure mathematician of the highest rank. It was not until the beginning of the twentieth century that research in pure mathematics, worthy of international repute, began once more to be produced in any considerable quantity. We shall have more to say of this shortly; in the meantime we continue with our story of Cambridge.

At both Oxford and Cambridge, the eighteenth century was a period of stagnation or even decay. In his autobiography Gibbon has left us a picture of one old university that will very well serve for both: idle students and still more reprehensible teachers, professors who never lectured, and some who never resided. Cambridge mathematical studies had gone the same way as the other

disciplines. The University degree examinations were still held in the medieval fashion, but these disputations had by now sunk to a mere farce.

Towards the end of the century, however, the first signs of a Cambridge revival appeared. A mathematical examination with some pretensions to seriousness came into being; although at first purely optional, it gradually ousted the old degree examination, so that in the early nineteenth century it became the sole test for the B.A. It was then that the examination acquired the name of the Mathematical Tripos. The results of the examination were published in order of merit; the first man on the list was called the Senior Wrangler; after him came the other Wranglers—these were the candidates who had been deemed worthy of a first class. Next in order came the Senior Optimes—these formed the second class. Finally, there were the third class men who were called, somewhat euphemistically, Junior Optimes; one suspects that by modern standards a fair number of these would have been refused a degree altogether. On Degree Day, when the successful candidates were presented to the Vice-Chancellor in the Senate House, a curious ceremony was observed. As the last Junior Optime—the bottom man on the list—came forward to be presented, from the public gallery there was lowered an enormous wooden spoon, which he received as a consolation prize. The expression "to get the wooden spoon" has since become proverbial.

As we have said, under the new arrangements every Cambridge man who wished to graduate had to go through the mathematical examination. At this stage of events, the ease with which one could scrape a third class showed a way out of the difficulty; even so, many men of ability were compelled to waste their time, and some had to leave the University without a degree. Even after another examination, the Classical Tripos, was instituted in 1822, the degree examination remained the same. And soon after that its standard was to be stiffened to a remarkable extent. That made things much worse for the nonmathematicians.

Quite early in the nineteenth century a handful of Cambridge men began to realize that it was high time to come out of the Great Sulk. In 1821 they formed what they called the Cambridge Analytical Society, whose aim was to familiarize British mathematicians with the work of the great Europeans. One of the leaders of this movement was Charles Babbage, who has since become famous as the father of the electronic computer. The task assumed by these young men was formidable indeed: for there is only one more conservative institution in the world than Cambridge, and that is Oxford. But it may be asserted that after ten years or so of propaganda their work began to show fruit. We thus enter the modern period of Cambridge mathematical studies, and with it the utterly fantastic story of the Mathematical Tripos in its golden days.

As we have already recounted, the Tripos was from its inception a competitive examination. By some process which has never been satisfactorily described, the fresh enthusiasm for mathematics which now burst upon the University transformed this examination into a high-speed marathon whose like has never been

seen before or since. It became far and away the most difficult mathematical test that the world has ever known, one to which no university of the present day can show any parallel. This is undoubtedly a sweeping statement; but the evidence for it is clear and overwhelming. The nineteenth century is, of course, the great period of Cambridge mathematical physics; it includes Ferrers, Green, Stokes, Kelvin, Clerk Maxwell, G. H. Darwin, Rayleigh, Larmor, J. J. Thomson—to mention only the top flight. Now all these men went through the Tripos mill; and it strained their abilities to the utmost.

At that time the teaching was entirely in the hands of the individual colleges, and much of it was grievously inadequate. But in any case, it could never have served the needs of the Tripos examination in its new form. Other methods of instruction were sought—and found. Any man coming to Cambridge and wishing to take a high place in the Tripos, at once put himself in the hands of a professional coach. The training was intensive and unremitting; it lasted for ten terms, with all the intervening long vacations as extra study periods. The examination was then taken early in the January of the undergraduate's fourth year of residence.

Each of the coaches divided his pupils into a number of small groups, who met him once or twice a week. At the meeting he would hand them back their solutions to the previous week's problems, and circulate in class his own set of solutions. While this was going on he would expound the new theorem or subject for study at the blackboard—which meant that a trainee had for part of the time at least to subdivide his attention between manuscript and lecture.

This relentless driving had two purposes before it: to train the candidate to the point where he could write out, with lightning speed and no hesitation, the proof of any theorem required by the syllabus; and where he could write out, at lightning speed and with almost no hesitation, the solution to any problem the examiners might set. Mere mathematical ability was not enough: rapidity of thought had to be added to it. Let me give some examples. Cayley, who afterwards occupied the Cambridge chair of pure mathematics for many years, took the Tripos in 1842. At that time a feature of the examination was a two-hour paper consisting entirely of problems; it was not expected that the candidates would do all the questions—they would naturally have a choice. On the evening after the problem paper, a friend called on Cayley to see how things were going; and, in the manner of friends, he brought bad news. "I've just seen Smith," he announced (Smith was a rival), "and do you know, he did all the questions within two hours." "Oh," remarked Cayley, "well, I cleaned up that paper in forty-five minutes."

Lord Kelvin, when he was just plain William Thomson of Peterhouse, was easily the best mathematician of his year, and was widely tipped for the Senior Wranglership. In fact, on the day that the results of the Tripos were published, he said to his college servant, "Oh, just go down to the Senate House, will you, and see who is Second Wrangler?" Soon after the man returned, and announced: "*You*, sir." Evidently there had been someone in the examination hall who could write, if not think, faster than Kelvin.

As may be imagined, coaching for the Tripos was a highly specialized profession; at any given time there were only two or three men of outstanding capacity for this odd perversion of learning. The most famous of all coaches was unquestionably E. J. Routh. Routh was a very considerable mathematician in his own right, and his books on dynamics are still standard works of reference. For a number of years he had a virtual monopoly on Senior Wranglerships, and many of the highest places were invariably taken by his pupils. Another celebrated coach was R. R. Webb of St. John's College. It may surprise most people to learn that the great mathematician W. H. Young also had a hand in the business: one hardly associates such goings-on with a man of his calibre. But it should be borne in mind that the business was very profitable.

In my student days at Cambridge, I attended some advanced lectures in differential geometry which were given by an elderly don named R. A. Herman. He was the last of the great coaches, a survival from a past epoch. Herman was a Fellow of Trinity College, of which Hardy and Littlewood also were members, and he had taught them both in his time, though their success could hardly be attributed to any of their instructors. All three of them had been Senior Wranglers. Nobody could recall when Herman had taken the Tripos—he was by then a legendary figure—but Hardy had triumphed in the year 1898 and Littlewood in 1908, the very last year of the old regime: the following year the new regulations came into force and with them the order of seniority disappeared. Herman had been a pioneer geometer in his day; it was he who introduced into England the use of the moving trihedron in differential geometry, and Hardy himself has put on record his gratitude to Herman for his inspiring lectures on this subject. But when I sat under him some thirty years later, the inspiration had all flickered out; he reminded one of an extinct volcano. Just occasionally, in a chance word or gesture, one could glimpse the man there had been.

So far we have dealt with the Tripos in its more superficial aspects. What was the examination itself like in those times? At the beginning, and indeed until the Analytical Society had brought about a change, the whole system remained tied to Newton. Out of blind loyalty to their Master, the examiners insisted as far as possible on maintaining a form and a substance of which he might have approved. Thus in problems concerning planetary motion or gravitational attraction, candidates were obliged to use the methods of classical geometry which Newton had employed in the *Principia* and which his own discoveries in the calculus had already rendered obsolete even before he composed the work. And in questions dealing with the calculus, the candidates had to adopt the bad notation of fluxions and fluants due to Newton, instead of the good notation of Leibniz which had long gained universal acceptance elsewhere.

After the reformers had had their way, such antiquated notions were discarded. All the same, the Tripos examination remained predominantly a test in applied mathematics. There was an excellent reason for this bias: the ex-

aminers themselves knew hardly any pure mathematics anyway. And the system was obviously self-perpetuating: with each generation the mantle descended from mathematical physicist to mathematical physicist. However, it should not for a moment be imagined that the examination was particularly concerned with applied mathematics in any serious sense of the term: the typical Tripos question, which has been parodied over and over again, was an unreal, often fantastically unreal, abstraction from the physical problem which had suggested it, whose sole object was to render it tractable to the candidates.

Moreover, the Tripos remained a highly conservative institution; with the passage of the years, the distinguished band of mathematicians, whom we have already named, continued to make fundamental contributions to the science; but the Tripos syllabus, generally speaking, kept a respectable distance behind them. Thus Bertrand Russell, who took the examination about 1890, has commented upon both the academic character of the courses and the time-lag in the curriculum. In his day, the Tripos examiners had not yet caught up with Maxwell's equations, which had been given to the world a generation earlier.

Clearly, then, another reformer was called for, one who would take the work on from the point where Babbage and his friends had left it. This might conceivably have been Cayley, who occupied the chair of pure mathematics from 1863 until his death in 1895. But Cayley, despite his eminence in the field of pure mathematics, was a professor of the old school, who looked upon the academic world of his time, saw that it was very bad, but continued to go his own way. Although Cayley lectured regularly, as he was strictly bound to do, upon various branches of pure mathematics, his audience was practically nonexistent. Often it consisted of no more than one pupil. But that pupil was Andrew Russell Forsyth.

This extraordinary man, whom I had the privilege to know in his capacity of Professor Emeritus at the Imperial College of Science, died in extreme old age in 1942; but his fame lives on. Everybody knows him as the author of the most successful book on differential equations that has ever appeared in any language; although it was first published as long ago as 1885, it is still being reprinted. I would venture the opinion that this work has done more than anything else to retard the true development of the subject; for over two generations it has continued to put wrong ideas into people's heads concerning the nature and scope of the theory and, thanks to the author's forceful and authoritative style, in this it has been overwhelmingly successful.

The truth is that Forsyth had the misfortune to be born a hundred years too late; in his mathematical outlook and technique, he was a man of the eighteenth century. His major work on the theory of differential equations, a colossal achievement in six volumes, is still today the only treatise in its class which is by a single hand; but a mere glance at the list of contents suffices to

reveal that, on the whole, Forsyth looks backward to Lagrange rather than forward to Cauchy. However, some knowledge of the rudiments of analysis was essential to an understanding of the work; and as the Cambridge men of his generation had none, the author, who had now succeeded to Cayley's chair, set himself the task of educating them. And this is where he comes into our story.

In 1893 there appeared the first edition of Forsyth's *Theory of Functions of a Complex Variable:* another production which cannot be described as anything less than colossal—even a German professor might have quailed before such a project. The book includes fairly complete accounts of the relevant work of Cauchy, Abel, Riemann, Weierstrass, Appell, and carries on the survey right up to the then contemporary researches of Klein and Poincaré. The style of the book is magisterial, Johnsonian; the author's powers of assimilation are well-nigh incredible—and yet, strange to say, despite his intentions and his absorption of the material, he never comes within reach of comprehending what modern analysis is really about: indeed whole tracts of the book read as though they had been written by Euler.

Nevertheless, for all its shortcomings, this was the work which brought modern pure mathematics into Cambridge. The young men at once began to imbibe it; and not the young men alone. My own copy of the book once belonged to R. R. Webb, the coach whom I have mentioned above, and from his pencilled notes in the margins it seems pretty clear that he was learning his function theory the hard way, much as any beginner would. The very fact that the book was written in the wrong spirit probably contributed to its great initial success. As Littlewood once put it, "Forsyth was not very good at delta and epsilon"; but neither was the public for whom he wrote: so author and readers met on common ground. In any case, it served as a stepping-stone to the real thing, which at that date was to be found only in French or German. Hardy has recorded that he himself first saw the light when he read the volumes of Jordan's *Cours d'Analyse*; and many other young men of his generation must have done likewise.

Within the space of ten years, Forsyth's treatise had achieved its aim. But it also accomplished something which its author had certainly never intended. For Cambridge now found itself equipped with a corps of modern pure mathematicians whose nominal leader was a living fossil firmly fixed in the Sadlerian chair. This grotesque situation seemed to all intents and purposes a permanent one: Forsyth's international reputation was enormous and in any case there was no possibility of removing him; it appeared as though he were there for life. But now fate took a hand in the game. In the year 1909 Forsyth, in the company of other scientists and their families, was travelling to a meeting of the British Association to be held in Canada. Among the party were the eminent physicist C. V. Boys and his wife Marion. Forsyth was then an apparently confirmed bachelor of 51; but he and Marion Boys fell in love with

one another. The end of it was that she decided to leave her husband; and this meant that Forsyth was compelled to resign his professorship, for in the Cambridge of those days there was no place for even the suggestion of divorce. It is pleasant to add that everyone concerned in this affair lived happily ever after; for it was generally conceded by all his acquaintances that the bereaved husband bore his loss with remarkable fortitude. (The former Mrs. Boys was a powerful personality.)

Forsyth survived his wife by many years; in fact he contrived to outlive everything—that was his tragedy. He had to retire from his chair at the Imperial College because he had reached the extreme age limit, although he commanded enough energy to have carried on for at least another five years. He set himself to learn Arabic and Persian; he wrote several enormous volumes on what were ostensibly branches of modern mathematics, all treated from the eighteenth century point of view; the Cambridge University Press, which made a fortune out of his earlier publications, must have lost a good deal of it on these. And all the time he was filling reams of paper with formulae and calculations; I happen to possess some manuscripts of his Cambridge lectures and also of some work on which he was engaged a year or two before his death: the differences between them, from the standpoint of calligraphy, are almost negligible.

I am pleased to relate that I have been able to pay one small tribute to this remarkable son of Cambridge. Some years ago, when I was asked to rewrite the article on Cayley for the *Encyclopaedia Britannica*, I took the opportunity to slip him in; and there, for some time to come, I hope he will stay.

But to continue with our main narrative. Although it almost goes without saying that Forsyth had himself been a Senior Wrangler (he had studied with Routh and Webb) and, moreover, was temperamentally inclined towards the Tripos kind of mathematics, yet he was one of the chief promoters of Tripos reform. At this point we may conveniently put a question which must have occurred to the reader very much earlier: how did such a fantastic sort of academic contest ever take root in the university? I think that the blame for it must be laid upon the system of college autonomy which we have previously described. In the absence of any strong central authority, the examination had fallen into the hands of private individuals, owing no responsibility to anyone; and until the university was able to regain some of its lost powers, there was little chance of breaking their hold over it. This reversion to the medieval form of government took place at about the turn of the century; the time was now ripe for reform.

As I have said, the new regulations for the Tripos came into force in 1909, the very year in which Forsyth, unbeknown to himself, was preparing to quit Cambridge. When I took the examination, nearly twenty years later, the net result of all the changes could be summarised as follows. In the first place,

the published order of merit had gone, and with it the rat race. Secondly, there was now a fair balance between pure and applied mathematics in both syllabus and examination questions. In the third place, a more advanced section of the old Tripos, which had been taken at a later stage by men who were seeking a fellowship, was incorporated in the undergraduate examination and entitled Schedule B, to distinguish it from the "elementary" part, known as Schedule A. (Incidentally, the University has since returned to the old practice: Schedule B, rechristened Part III, is now generally taken after a fourth year of residence.)

What was the new examination like? No doubt, if any of the high Wranglers from the past century could have returned to scrutinise our papers, they would have pronounced them pretty easy. To most of us, however, they appeared quite otherwise, particularly at first sight, in the examination room. To begin with, although the notorious Tripos trickery had now given place to a more mature outlook, there was still sufficient need for ingenuity to make the whole proceeding a highly risky business. The typical Schedule A question was a three-decker: first the candidate would be asked to prove a theorem; then would come a problem based more or less on this theorem; and thirdly, another problem even less based than the first. In fact, despite all appearances to the contrary, this last might break fresh ground: that was the sting in the tail. Everybody knew that only complete answers to questions really counted, and that the postscript usually mattered more than the rest. Hence a certain general foreboding. A candidate, even a well-prepared one, might go into the examination on the Monday morning and find himself unable to do a single complete question; if, unduly depressed by this failure, he had the same experience on Monday afternoon, then it was all over save the post-mortem.

The kind of question one had to face may be illustrated by the following example—possibly fictitious, for I do not remember seeing it anywhere. In the first part the candidate is asked to obtain the general solution to Laplace's equation in three dimensions, in terms of Legendre functions. Next, he is required to apply his results to the problem of a conducting sphere in a given field. In the third part the sphere is replaced by an ellipsoid which is nearly spherical. Now the fun begins: the answer to this part of the question is stated—not as a guide to the solver, but in order to make the question more difficult. For as the desired result is merely an approximation, anybody could produce an answer of sorts: the real difficulty is to arrive at the formula stated by the examiners. In questions of this type one might polish off the first two parts in no time at all, only to waste up to an hour on the third. And that way madness lies.

Schedule B was a quite different affair, not without its own peculiar troubles. Whereas Schedule A was taken by everyone, and on it one's class was usually decided, Schedule B was optional, that is to say, if any undergraduate possessed the necessary tenaciousness, cunning, and indiscipline, he

might be able to persuade his tutor to let him out of it. Assuming this was impracticable, he found himself confronted by a dilemma. Schedule B was based on the advanced courses in pure and applied mathematics; the questions set in it were of the longish essay type, each taking about an hour to write out. Once again, only complete answers to the questions were really of much use. Now the courses were numerous and comprehensive, while the total number of questions set was comparatively small. If a candidate chose to take a few courses only, he might find on the fatal day that he was unable to answer a single question on a particular paper. I vividly recall one session of our examination at which, ten minutes after the papers had been given out, a candidate rose in his place and walked slowly out of the hall. This act did nothing to cheer those of us who remained.

One might grasp the other horn of the dilemma by electing to take a great number of courses so as to insure against this disaster. But then there arose another serious difficulty: how could one commit all this material to memory? Here again, a fragmentary knowledge of the syllabus was only of doubtful value; for by sheer bad luck the questions set might weave in and out of the candidate's recollections and so lead inevitably to incomplete answers.

Luck certainly played a considerable part in the examination. I myself had some of each sort, though admittedly more good than bad. I took a term's course of lectures given by Littlewood on the foundations of function theory—this course, or a modified form of it, is still being reprinted as a paperback. But the printed version can give no idea of how delightful the lectures actually were: for Littlewood is one of the wittiest mathematicians that Cambridge, or indeed any other university, has ever produced. When, however, we came to the examination I found to my dismay that he had set about the most difficult question in the course. This was it: "Prove that, if a and b are any two given numbers, then one of the following possibilities must hold: either a is less than b, or a is equal to b, or a is greater than b." Perhaps this result may seem obvious to some of my readers; it certainly seemed obvious to me at the time—in the sense in which it appears obvious to them. But I distinctly recalled that, during the lectures, Littlewood had made a fearful mess of the demonstration; it had taken him the best part of an hour to write up on the blackboard. And I had the feeling that, on the present occasion, he would settle for nothing less. So, apart from the jokes, that course had to be written off as a dead loss.

One day my director of studies said to me: "I see that Mr. Pars" (they were all Misters in those days—no Cambridge man would have been seen dead with a Ph.D.) "is giving a course on general dynamics. I think you might attend." Actually there was no subject I cared less about; but this was a command, and so I attended. Now L. A. Pars was one of those lightning performers in whom Cambridge has always specialised. It is true that he wrote every single word upon the blackboard, but at such a pace that it was next to impos-

sible to keep up with him: to understand what he was doing was quite out of the question. After two terms of this sort of treatment I had a wad of notes as bulky as Whittaker's treatise, on which the lectures were mainly based. An important difference between my account of the subject and Whittaker's was that he knew what he was writing about. As events showed, however, this was irrelevant; for in the examination I encountered two questions on the course which, although ostensibly devoted to problems of dynamics, were really the purest of pure mathematics. I managed to do them both.

This little incident brings us to the grievance which many students nourished against the Tripos system as a whole. In a reasonably balanced examination for the degree a candidate whose interests lay in pure mathematics would have been required to take all the pure mathematics courses together with a selection of the applied; and similarly for a specialist in applied mathematics. But the Cambridge plan insisted on the double dose; and for all but the highly gifted minority this laid an almost intolerable burden upon the conscientious student, to say nothing of the fact that it took no account of the use to which he would turn his knowledge in after life. Experience shows that a taste for one main branch or other of the subject is as a rule acquired fairly early, and that a change is seldom made after. Consequently most of us felt in our bones that we were wasting an awful lot of our time.

And so we arrive at the examination itself. This too could have been arranged better, one thought. Schedule A consisted of six papers: Monday, Tuesday, Wednesday, from 9 A.M. to noon, and then from 1:30 to 4:30 P.M. On the following Monday, Tuesday and Wednesday we sat for Schedule B, which likewise consisted of six papers, each of three hours duration. What went on during those thirty-six hours I cannot now recall in any detail; all I can say is that it was a kind of continuous nightmare. But there was an opening episode which still sticks in my memory. Our examination was held in the great hall of King's College, under the shadow of the famous Gothic chapel; so those candidates who were Scholars of King's had only to cross the college court. I happened to be acquainted with one of them—he was an eccentric individual hailing from Lancashire, and he was notorious for the fact that during the vacations he used to earn a living by playing the organ at the cinema in his home town. At that time vacation work was almost unheard of in England, and his conduct was generally regarded as very queer and perhaps slightly scandalous. As I have said, he had only to cross the court to reach the examination hall. He was wearing his gown, as was necessary at lectures and examinations, but he hadn't bothered to change his slippers. So the officials on duty declined to admit him because he was improperly dressed (I happen to know this as I was sitting near the door and overheard the conversation). Whether he subsequently returned, correctly shod for the occasion, or whether he remained shut out forever I really cannot say; for only a few minutes later I had many other things to worry about.

I remember clearly enough the closing scene of our Tripos. As soon as the last papers were handed in, two fellow-sufferers—myself and a friend—rushed from the hall and walked as fast as we knew how the three miles down to the river where the first of the May races were due to begin. There, in the midst of a crowd of undergraduates and their guests, we soon banished the Tripos from our thoughts. I secretly vowed never to take another examination in my life; and this vow, I am glad to say, I have kept.

* * *

Here my own reminiscences of the Tripos come to an end; but of course the story goes on. Various generations of Cambridge men have each shaped the examination according to their light, but the work is never complete and probably never will be. Other intending reformers of the Tripos are even now waiting in the wings; indeed some among them would reform it altogether. Such a notion, startling as it may appear, is by no means novel; it was held more than forty years ago, by Hardy himself, who had backed the 1909 reform as only a first stage of the program. Hardy firmly believed that the Tripos was an unmitigated evil, for which one must blame the inferior performance of British pure mathematicians *vis-à-vis* their European colleagues. So, away with the examination.

Now the weakness of this argument resides in its lack of supporting evidence. It would be very difficult to unearth any specific cases of careers which have undoubtedly been ruined or even seriously damaged by the Cambridge mathematical system: on the other hand, the supporters of the status quo can for their part point to a long line of distinguished mathematical physicists, some of whom we have already mentioned, who achieved success either because or in spite of it: looking at their record one could scarcely suppose that they would have done more or better work had they been spared the ordeal of the examination; and, in the past, what an ordeal it was!

Abolitionists are such charming people; their motives are so patently pure, and only rarely do they foresee the full consequence of their projected panaceas. All his life long Hardy moved in the highest academic circles and tutored the most talented of young men. Had he troubled to consult any lecturer from a provincial university, or even (it may be) a don from a Cambridge college less exalted than Trinity, he could easily have learned a simple but significant truth: if students know beforehand that a particular subject is not to be examined upon, they will, almost to a man—or a woman—altogether decline to study it. Even Forsyth could have told Hardy that much: for he had been a professor at London University.

ON TWO CLASSICAL THEOREMS OF ALGEBRAIC TOPOLOGY

W. M. BOOTHBY, Washington University, St. Louis, Missouri

Among the most beautiful and frequently quoted results of elementary algebraic topology are the following two theorems, the first due to L. E. J. Brouwer and the second to Brouwer and H. Poincaré.

THEOREM I. (Brouwer Fixed Point Theorem) *Each continuous map of the solid unit ball into itself has a fixed point.*

THEOREM II. *Any continuous field of vectors tangent to an even dimensional sphere must be zero at some point.*

For theorems which have stimulated so much further research, beginning with the work of H. Hopf and continuing to the present, and whose content is so clear and easy to state, they are surprisingly difficult to prove, even in the simplest cases—the unit disk and the ordinary 2-sphere. Proofs are customarily given in standard courses in algebraic topology, but only after a fairly extensive theory is developed. In a brief, but very readable and elegant book [3], J. Milnor gave relatively simple proofs, based in part on a very original approach due to M. Hirsch [4]. Since this book goes into many generalizations of these theorems, it introduces and uses some techniques of differential topology which we wish to avoid, for example Sard's theorem. In this paper it is our purpose to make use of the fact that several advanced calculus texts discuss both manifolds and exterior differential forms in the presentation of multiple integrals and Stokes' theorem, and thus to rewrite the standard classical proofs using these techniques to avoid overt use of algebraic topology. In fact the present treatment goes very little beyond material to be found in the texts of either Fleming [6] or Devinatz [7] or in the more specialized book of Flanders [5], and the proofs of these theorems become, then, a somewhat delicate exercise in advanced calculus.

1. Preliminaries. Although we suppose the reader to be acquainted with those portions of the advanced calculus texts mentioned above which deal with manifolds and maps, differential forms, Stokes' theorem, and the Poincaré lemma, we shall nevertheless briefly discuss some of these topics by way of review. The manifolds—always supposed differentiable—which we shall need are particularly simple ones: Euclidean n-space R^n represented as n-tuples of real numbers—hence furnished with a fixed coordinate system covering the entire space; its submanifold S^{n-1} (the $n-1$ dimensional sphere con-

Prof. Boothby received his Michigan Ph.D. under W. Kaplan. Before his present position at Washington University, he was on the staff of Northwestern University, and he spent a post-doctoral fellowship of the American Swiss Foundation at the E.T.H., Zurich. Since then he has held an NSF fellowship at the I.A.S., Princeton, and a sabbatical year at the University of Geneva, supported both by the NSF and the Amer. Swiss Found. His main research is in differential geometry and Lie groups. *Editor.*

sisting of all $(x_1, \cdots, x_n)$ with $\sum x_i^2=1$); and the "cylinder" $R\times S^{n-1}$, which is an n-dimensional submanifold of $R^{n+1}=R\times R^n$. These manifolds are *orientable*, which means that they possess a covering by coordinate neighborhoods with the property that whenever two of these neighborhoods overlap, the differentiable functions giving one set of coordinates in terms of the other have positive Jacobian. If a connected manifold is orientable, then there are two disjoint classes of such coverings; a choice of one of these classes is said to *orient* the manifold. Note that the space R^n is automatically oriented by the fixed coordinate system associated with it. In the sequel *manifold* will always mean connected, orientable, and differentiable manifold.

Throughout this paper we shall use the term *differentiable* to mean C^∞ (indefinitely differentiable). Thus a real valued function $f(x_1, \cdots, x_n)$ on an open subset $U\subset R^n$ is differentiable if it has continuous partial derivatives of all orders at every point of U; manifolds are assumed to be covered by coordinate neighborhoods for which change of coordinates on overlapping neighborhoods is C^∞; and mappings from one manifold to another, unless otherwise stated, will have expressions in local coordinates which are indefinitely differentiable. In particular, any real-valued function on an open subset U of a manifold M is differentiable if its restriction to each local coordinate neighborhood on M is C^∞ when expressed in terms of the local coordinates, i.e., as a function on an open subset of $R^n (n=\dim M)$.

These definitions should be fairly familiar, but less familiar is the definition of a differentiable function f on an *arbitrary* subset A of a manifold M. In this case we say that f is differentiable if it can be extended to a differentiable function f^* defined on an open subset U such that $A\subset U\subset M$: this reduces it to the familiar case above. But there is a difficulty; even a trivial case —A consisting of a single point of R^n—shows that the *values* of the derivatives may depend on the extension and thus have no meaning for the function f itself. To avoid this problem we restrict the arbitrariness of A: we shall say that a connected subset A of a manifold M is a *domain* if it is the closure of its interior, i.e., $A=\bar{V}$, where $V=\operatorname{Int} A$ is an open subset of M. If f is a differentiable function on a domain A, then any two extensions which are differentiable on an open set U containing A must have identical derivatives of all orders on the open set $V\subset U$ and thus on $\bar{V}=A$. It follows that the derivatives of all orders are well-defined, continuous functions on A. In particular, the restrictions of local coordinates of M to a domain A are differentiable functions on A, hence coordinates on A. Therefore the same covering which orients M orients any domain A of M. If $A\subset M$ and $B\subset N$ are domains of manifolds M and N, then differentiability of a map $F: A\to B$ is defined as for manifolds, i.e., in terms of local coordinates.

Many of the domains we use are even more restricted. A *regular domain* D of a manifold M is a domain which is compact, connected, and whose boundary, denoted ∂D, is an $n-1$ dimensional submanifold which divides M into two components. Examples are: (i) the unit interval $I=\{x\in R \mid 0\leqq x\leqq 1\}$ in

the manifold $M=R$, (ii) the unit ball $D^n=\{(x_1, \cdots, x_n)\in R^n \mid \sum x_i^2=1\}$ in the manifold $M=R^n$, and (iii) M itself if M is compact—in this case ∂M is empty. Regular domains are the domains of integration in the treatment of Stokes' theorem found in [6]; it is important for this theorem that D is oriented by the restrictions to D of the covering by coordinate neighborhoods which orients M—moreover, this orientation of D induces an orientation on ∂D. Thus the fixed orientation of R^n by the single system of coordinates which covers it also orients D^n and its boundary, the unit sphere S^{n-1}. Not all the domains which we consider are regular; it will also be necessary for us to use spaces which consist of the cartesian product $I\times D$ of the unit interval and a regular domain D of a manifold M, for example, the solid cylinder $I\times D^2$. Spaces of this type are domains of $R\times M$, a manifold; they differ from regular domains only in that their boundary is not quite a submanifold. Their use presents no new difficulties.

Since differentiable functions, coordinates, and maps have meaning for a domain A of a manifold M, we may also consider differential forms and their exterior derivatives on A just as we do on a manifold. The forms of degree q are denoted by $\wedge^q(A)$ and the totality of all forms of all degrees by $\wedge(A)$. This set forms an associative algebra over the real numbers (products are denoted with a wedge $\wedge$) and each $\wedge^q(A)$ is a linear subspace. The 0-forms are just the differentiable functions on A, and $\wedge^q(A)$ for $q>n=\dim M$ contains only the 0-form (0 of the algebra). The exterior derivative d is a *linear* map whose iterate $d^2=0$ and which takes each q-form to a $q+1$ form. In particular, it assigns to each function (0-form) its differential and to each n-form the only possible $n+1$ form, namely 0. If $\alpha\in\wedge^r(A)$ and $\beta\in\wedge^s(A)$, we have

$$\alpha\wedge\beta = (-1)^{r+s}\beta\wedge\alpha \quad \text{and} \quad d(\alpha\wedge\beta) = d\alpha\wedge\beta + (-1)^r\alpha\wedge d\beta.$$

Thus $\wedge(A)$ is not commutative, nor is d a homomorphism. However, these properties do imply that the kernel of the linear map d, i.e., those forms ω such that $d\omega=0$, is a subalgebra and that the image of d, i.e., those forms θ such that $\theta=d\eta$ for some η, is an ideal within it (forms in the kernel are called *closed* and those in the image *exact*). Finally we mention the fact that any differentiable map $F:A\to B$ of domains induces an algebra homomorphism $F^*:\wedge(B)\to\wedge(A)$ in the opposite direction. It has the properties (i) $F^*\circ d=d\circ F^*$, (ii) under composition of maps: $(F\circ G)^*=G^*\circ F^*$, and (iii) the identity map induces the identity isomorphism.

A form ω on A is completely determined by its expressions in the local coordinates, which we now discuss briefly. In local coordinates or on an open subset of R^n, the differentials $dx_1, \cdots, dx_n$ of the coordinate functions form a set of generators for the algebra, and $\omega\in\wedge^q(U)$ may be written

$$\omega = \sum a_{i_1\cdots i_q} dx_{i_1}\wedge\cdots\wedge dx_{i_q} \qquad \text{with } dx_i\wedge dx_j = -\,dx_j\wedge dx_i.$$

Then the properties mentioned above imply that

$$d\omega = \sum da_{i_1\cdots i_q} \wedge dx_{i_1} \wedge \cdots \wedge dx_{i_q} = \sum \frac{\partial a_{i_1\cdots i_q}}{\partial x_j} dx_j \wedge dx_{i_1} \wedge \cdots \wedge dx_{i_q}.$$

They also imply that whenever a map $F: A \to B$, A and B domains of manifolds of dimensions m and n respectively, is expressed in local coordinates, say with $y^\alpha = y^\alpha(x_1, \cdots, x_n)$, $\alpha = 1, \cdots, m$, then $F^*: \wedge^q(B) \to \wedge^q(A)$ is computed in these coordinates by replacing each dy_i by its corresponding expression as a differential of a function of $x_1, \cdots, x_n$ and regarding the coefficients as composite functions of the x-coordinates through the y's. For example, if $q = 1$, then

$$F^*\left[\sum b_i(y) dy_i\right] = \sum_i \sum_j b_i(y(x)) \frac{\partial y_i}{\partial x_j} dx_j.$$

An important special case of the map F^* induced by a differentiable map F is the inclusion (or identity) map J of a subset into the set. Examples of inclusion we need are a domain $A \subset M$ into M, a submanifold N of M into N (especially, for us, S^{n-1} into R^n), and of a coordinate neighborhood U of M into the manifold M. In these cases we usually adopt the notation ω_A, ω_N, ω_U, etc., for the image $J^*\omega$ of a form ω on the ambient manifold M. The form thus obtained is called the *restriction* of the original form to the subset and it will be used below to obtain forms on both D^n and S^{n-1} from forms on R^n, or even, in the S^{n-1} case, from forms on D^n, since S^{n-1} is a submanifold of D^n as well as of R^n. In computation with forms, one makes repeated use of the fact that a form on $A \subset M$ is determined by its restrictions to each neighborhood of a covering of coordinate neighborhoods of M. In this case, as already seen, we allow ourselves to use ω, rather than ω_U, ω_V, etc., for these restrictions.

2. One-parameter families of differential forms. Let D be a domain of a manifold M, possibly all of M. We need to consider a family of q-forms ω_t on D parametrized by the variable t with $0 \leqq t \leqq 1$. In local coordinates $x_1, \cdots, x_n$ of a coordinate neighborhood U of M the forms would be written

$$\omega_t = \sum b_{i_1\cdots i_q} dx_{i_1} \wedge \cdots \wedge dx_{i_q},$$

the coefficients $b_{i_1\cdots i_q}$ being differentiable functions of t, $x_1, \cdots, x_n$. It is natural in this context to consider the domain $I \times D$ in the manifold $R \times M$. Each coordinate neighborhood U of M determines a coordinate neighborhood $R \times U$ of $R \times M$ with local coordinates t, $x_1, \cdots, x_n$, where $x_1, \cdots, x_n$ are the local coordinates in U. We use only coordinates of this kind on $R \times M$ and its domain $I \times D$. Let t be a fixed element of the unit interval I and define the map $J_t: D \to I \times D$ by $J_t(p) = (t, p)$ and let $\tilde{\omega}$ be a q-form on $I \times D$. Then $\omega_t = J_t^* \tilde{\omega}$ defines, for each $0 \leqq t \leqq 1$, a q-form on D. The collection ω_t for all $t \in I$ is a one-parameter family of q-forms on D. Conversely, any one-parameter family ω_t on D, given in local coordinates as described above, determines an $\tilde{\omega}$ on $I \times D$ of which it is the image under the maps J_t. The latter definition

is more precise since it does not involve local coordinates and thus relieves us of the necessity of checking independence of coordinates.

More generally, consider an arbitrary q-form $\tilde{\omega}$ on $I\times D$. Using local coordinates as described above, its expression on $I\times U$ will have the form

$$(*)\quad \tilde{\omega} = \sum a_{i_1\cdots i_{q-1}} dt \wedge dx_{i_1} \wedge \cdots \wedge dx_{i_{q-1}} + \sum b_{k_1\cdots k_q} dx_{k_1} \wedge \cdots \wedge dx_{k_q}.$$

In general, the coefficients depend on t as well as $x_1, \cdots, x_n$. In fact t, the first entry of the pair $(t, p)\in \boldsymbol{R}\times \boldsymbol{M}$, is a real-valued function defined on *all* of the manifold $\boldsymbol{R}\times \boldsymbol{M}$; hence its differential dt is a globally defined 1-form. It follows that each q-form $\tilde{\omega}$ on $I\times D$ splits into the sum of a form which contains dt as factor and one which does not—the decomposition above does not depend on local coordinates. Note in passing that because t is constant in the definition of J_t, $J_t^*(dt)=0$ and more generally

$$J_t^*(dt \wedge \theta) = J_t^*(dt) \wedge J_t^*(\theta) = 0.$$

We use these remarks to define a linear operator $\mathcal{J}: \wedge^q(I\times D)\to \wedge^{q-1}(D)$ which plays a crucial role in the theory. Let $\tilde{\omega}$ be a q-form on $I\times D$ given in local coordinates by the expression (*) above, then

$$\mathcal{J}\tilde{\omega} = \sum \left[\int_0^1 a_{i_1\cdots i_{q-1}} dt\right] dx_{i_1} \wedge \cdots \wedge dx_{i_{q-1}}.$$

Since only the coordinates $x_1, \cdots, x_n$ of the coordinate neighborhood U of M are involved on the right, this gives a form on D in local coordinates. (Note that it is zero if dt is not a factor of $\tilde{\omega}$.) The verification that this mapping is independent of local coordinates is straightforward and may be found in the references cited, so we do not give it here. The most important property of the operator $\mathcal{J}$ is embodied in the following lemma, in which ω_0, ω_1 denote $J_0^*\tilde{\omega}$, $J_1^*\tilde{\omega}$ respectively:

LEMMA 1. *Let D be a domain of a manifold M (possibly all of M), $\tilde{\omega}$ be a q-form on $I\times D$, and $\mathcal{J}\tilde{\omega}$ the $(q-1)$-form on D defined above. Then we have* $\mathcal{J}d\tilde{\omega}+d\mathcal{J}\tilde{\omega}=\omega_1-\omega_0$.

The proof, reduced to local coordinates, is an application of Leibniz rule for differentiating under the integral together with the elementary properties of differential forms summarized in the previous section. It is proved in detail in the references, e.g., p. 27 ff. of [5] or formula (7–20) on p. 280 of [6], and we do not repeat it here. The present statement is somewhat more general in that D may be either a manifold, including an open subset of a manifold—which is also a manifold—or a more general domain of the type discussed in the previous section in connection with differentiability. The formal details are not changed by this increase in generality. Just as in the references [5], [6], and [7], we give as first application Poincaré's lemma, but in our case for the closed unit ball D^n as well as the open unit ball.

COROLLARY (Poincaré's Lemma). *For $q>0$ each closed q-form on the closed ball $D^n \subset R^n$ is exact. The same statement holds for its interior, the open ball B^n.*

Proof. Let θ be a q-form on D^n with $d\theta=0$. We define $H: I\times D^n \to D^n$ by $H(t, x_1, \cdots, x_n)=(tx_1, \cdots, tx_n)$; then $H(0, x_1, \cdots, x_n)=(0, \cdots, 0)$ and $H(1, x_1, \cdots, x_n)=(x_1, \cdots, x_n)$. If we let $\tilde{\theta}=H^*\theta$ and adopt our earlier notation: $(H^*\theta)_t=J_t^*\tilde{\theta}$ for the 1-parameter family of forms defined by $\tilde{\theta}$ on D^n, then we see that $(H^*\theta)_0=0$ and $(H^*\theta)_1=\theta$. Since $dH^*\theta=H^*d\theta=0$, we have $d\tilde{\theta}=0$. By Lemma 1, $d\vartheta\tilde{\theta}=\theta$; it follows that θ is exact as claimed.

REMARK. From the proof it is clear that the theorem holds for any star-shaped open set or domain of R^n.

3. A differentiable version of the theorems. In this section we apply the preceding lemma, Stokes' theorem, and the facts mentioned earlier about the behavior of differential forms with respect to mappings in order to establish C^∞ versions of Theorems I and II. The following lemma is used in both proofs:

LEMMA 2. *Let Ω be the $(n-1)$-form on S^{n-1} obtained by restricting to S^{n-1} the form*

$$\tilde{\Omega} = \sum_{j=1}^{n} (-1)^{j-1} x_j dx_1 \wedge \cdots \wedge dx_{j-1} \wedge dx_{j+1} \wedge \cdots \wedge dx_n$$

defined on all of R^n. Then Ω is closed but not exact.

Proof. The form Ω is closed since $d\Omega$ has degree n which is greater than the dimension of S^{n-1}, thus $d\Omega=0$. If we suppose there exists an $(n-2)$-form θ on S^{n-1} such that $\Omega=d\theta$, then applying Stokes' theorem, we obtain

$$\int_{S^{n-1}} \Omega = \int_{S^{n-1}} d\theta = \int_{\partial S^{n-1}} \theta = 0,$$

because ∂S^{n-1} is empty. However, $S^{n-1}=\partial D^n$, and another application of Stokes' theorem shows that

$$\int_{S^{n-1}} \Omega = \int_{\partial D^n} \tilde{\Omega} = \int_{D^n} d\tilde{\Omega} = n \int_{D^n} dx_1 \wedge \cdots \wedge dx_n,$$

which is clearly not zero. Thus our assumption that Ω is exact leads to a contradiction.

The next lemma is the standard one used in the classical proof. In its continuous version it is of interest in its own right; it says that a solid ball cannot be retracted onto its boundary.

LEMMA 3. *There is no differentiable map $F: D^n \to S^{n-1}$ such that on the boundary F is the identity, i.e., such that $F(\mathbf{x})=\mathbf{x}$ for every $\mathbf{x}\in S^{n-1}$.*

Proof. We suppose that such a map exists and obtain a contradiction. Let Ω be the form on S^{n-1} defined in Lemma 2 and let $J: S^{n-1} \to D^n$ be the inclusion map. Then $F \circ J: S^{n-1} \to S^{n-1}$ is just the identity map of S^{n-1}. Thus $(F \circ J)^* = J^* \circ F^*$, implies $J^*(F^*\Omega) = \Omega$. Now $F^*\Omega$ is a form on D^n and since $dF^*\Omega = F^*d\Omega = 0$, it is closed. According to the Corollary to Lemma 1 it is exact, i.e., we have $F^*\Omega = d\theta$ for some form θ on D^n. This means that $\Omega = J^*(F^*\Omega) = J^*d\theta = dJ^*\theta$ so that Ω is also exact. But we have proved in Lemma 1 that this is not the case. It follows that no such map F exists.

From this lemma we now obtain by the usual method ([3] or [8]) the differentiable version of Theorem I.

THEOREM I′. *If $G: D^n \to D^n$ is a differentiable map, then there is an $\mathbf{x} \in D^n$ such that $G(\mathbf{x}) = \mathbf{x}$.*

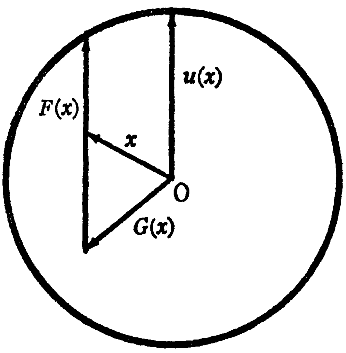


FIG. 1.

Proof. If there is no such point, then $\mathbf{x}$ and $G(\mathbf{x})$ determine a line; let $F(\mathbf{x})$ be its point of intersection with S^{n-1} as we move along the line in the direction $G(\mathbf{x})$ to $\mathbf{x}$. For $\mathbf{x} \in S^{n-1}$ this implies $F(\mathbf{x}) = \mathbf{x}$; so that $F: D^n \to S^{n-1}$ and F is the identity on S^{n-1}. Moreover, treating $\mathbf{x}$, $G(\mathbf{x})$, etc., as vectors, $F(\mathbf{x})$ is given by [3]

$$F(\mathbf{x}) = \mathbf{x} + \lambda(\mathbf{x})u(\mathbf{x}),$$

where $u(\mathbf{x})$ is the unit vector $(\mathbf{x} - G(\mathbf{x}))/\|\mathbf{x} - G(\mathbf{x})\|$ and $\lambda(\mathbf{x})$ is the scalar defined by $\lambda(\mathbf{x}) = -\mathbf{x} \cdot u + [1 - \mathbf{x} \cdot \mathbf{x} + (\mathbf{x} \cdot u)^2]^{1/2}$; the dot indicates inner product. It is easily checked from the geometry (Fig. 1) that the term in brackets is strictly positive so that $\lambda(\mathbf{x})$ is C^∞. Thus $F: D^n \to S^{n-1}$ is a differentiable map which is the identity on S^{n-1}, contradicting Lemma 3. Hence no such $G(\mathbf{x})$ exists.

We conclude this section by using Lemma 2 to give the differentiable version of Theorem II.

THEOREM II′. *There is no smooth, nowhere vanishing field of vectors tangent to an even dimensional sphere.*

Proof. Again following the classical proof of Alexandroff-Hopf [8], we first note that if the unit sphere S^{n-1} has a nonvanishing, differentiable, tangent vector field, then it has a differentiable field of unit tangent vectors: simply divide each vector by its length. Given a field of unit tangent vectors $u(x)$ to S^{n-1}, with $u(x)$ tangent at x, we have $x\cdot x=1$, $u\cdot x=0$, and $u\cdot u=1$. The map $H:I\times S^{n-1}\to S^{n-1}$ given by the formula (in vector notation)

$$H(t, x) = (\sin \pi t)u - (\cos \pi t)x$$

coincides with the identity map $J:x\to x$ when $t=1$ and with the antipodal map $A:x\to -x$ (which takes each $x\in S^{n-1}$ to the point at the opposite end of its diameter) when $t=0$. Let $\tilde{A}:R^n\to R^n$ be the reflection in the origin, i.e., $\tilde{A}:x\to -x$. If $\tilde{\Omega}$ is the $(n-1)$-form of Lemma 1, then $\tilde{A}^*\tilde{\Omega}=(-1)^n\tilde{\Omega}$. Let $J:S^{n-1}\to R^n$ be the inclusion map; we have defined $\Omega=J^*\tilde{\Omega}$. Obviously $J\circ A=\tilde{A}\circ J$ so that $A^*\Omega=A^*\circ J^*\tilde{\Omega}=(-1)^nJ^*\tilde{\Omega}$ and $A^*\Omega=(-1)^n\Omega$.

Now consider the form $H^*\Omega$ on $I\times S^{n-1}$; we have $dH^*\Omega=H^*d\Omega=0$, $(H^*\Omega)_1=\Omega$, and $(H^*\Omega)_0=A^*\Omega$. We assume n is odd, which means that $A^*\Omega=-\Omega$. Combining these facts and applying Lemma 1, we have

$$d\mathcal{J}(H^*\Omega) = \Omega - A^*\Omega = 2\Omega.$$

But this means that Ω is exact: $\Omega=d(\frac{1}{2}\mathcal{J}H^*\Omega)$, contrary to Lemma 2.

REMARK. When n is even, say $n=2m$, there always exists a differentiable, nowhere-vanishing vector field tangent to S^{n-1}. In fact the vector u at $x=(x_1, x_2, \cdots, x_{2m})$, $x\in S^{n-1}$, whose components are $(x_2, -x_1, x_4, -x_3, \cdots, x_{2m}, -x_{2m-1})$ is a unit vector with $u\cdot x=0$; hence it is tangent at x to S^{n-1}. We shall discuss some further cases in Section 5.

4. The proof of Theorems I and II. For the continuous case of Theorems I and II, one needs an approximation theorem. This may be proved directly (see Appendix) or by appeal to the Weierstrass Approximation Theorem, which is proved in [7] but only for the case of an interval on the real line.

APPROXIMATION LEMMA. *Let K be a compact subset of R^n and f a continuous, real-valued function defined on K. Then for any $\epsilon>0$ there is a C^∞ function $g(x)$ on R^n such that $|f(x)-g(x)|<\epsilon$ for all $x\in K$.*

In fact, according to the Weierstrass Approximation Theorem [9, p. 133] there is a *polynomial* $g(x)$ with the required property. For our purpose any C^∞ function $g(x)$ will suffice, and we are only interested in the case $K=D^n$. An easy, standard proof of the lemma in just the generality we require is given in an appendix. Using this lemma (for $K=D^n$) we derive from Theorems I′ and II′, the Theorems I and II of the introduction.

Proof of Theorem I. As in [3] suppose that $F:D^n\to D^n$ is a continuous map without fixed points. Using compactness we choose $\epsilon>0$ so that $\|F(x)-x\|>3\epsilon$ on D^n. Let $G_0(x)$ be a differentiable ϵ-approximation to $F(x)$ on D^n; the existence of $G_0(x)$ may be seen by applying the preceding lemma to each of the co-

ordinate functions of $F(\mathrm{x})=(f_1(x),\cdots,f_n(x))$. If $\mathrm{x}\in D^n$, then $G_0(\mathrm{x})$ may be outside D^n, but we must have $\|G_0(\mathrm{x})\|<1+\epsilon$. Thus $G(\mathrm{x})=1/(1+\epsilon)G_0(\mathrm{x})$ is a differentiable map of D^n into D^n, and for $\mathrm{x}\in D^n$ it must satisfy

$$\begin{aligned}\|F(\mathrm{x})-G(\mathrm{x})\| &= \left\|F(\mathrm{x})-\frac{1}{1+\epsilon}G_0(\mathrm{x})\right\| = \frac{1}{1+\epsilon}\|\epsilon F(\mathrm{x})+F(\mathrm{x})-G_0(\mathrm{x})\| \\ &\leqq \epsilon\|F(\mathrm{x})\|+\|F(\mathrm{x})-G_0(\mathrm{x})\| < 2\epsilon.\end{aligned}$$

By Theorem I′ we see that $G(\mathrm{x})$ must have a fixed point. However we have the following inequalities:

$$\begin{aligned}\|G(\mathrm{x})-\mathrm{x}\| &= \|(F(\mathrm{x})-\mathrm{x})-(F(\mathrm{x})-G(\mathrm{x}))\| \\ &\geqq \|F(\mathrm{x})-\mathrm{x}\|-\|F(\mathrm{x})-G(\mathrm{x})\| \\ &\geqq 3\epsilon-2\epsilon=\epsilon.\end{aligned}$$

This is an obvious contradiction, hence no such F can exist.

Proof of Theorem II. We consider the unit sphere S^{n-1} with $n-1$ even and suppose that $u(\mathrm{x})$ is a continuous field of tangent vectors which is never zero. Thus, dividing each vector by its length, we may assume in fact that for each $\mathrm{x}\in S^{n-1}$ the tangent vector $u(\mathrm{x})$ has unit length, i.e., we may assume $\mathrm{x}\cdot u\equiv 0$, and $\|u\|\equiv 1\equiv\|\mathrm{x}\|$. For $\mathrm{x}=(x_1,\cdots,x_n)\in R^n$ we let $r^2=\sum x_i^2=\|\mathrm{x}\|^2$; then we may extend the component functions $u_1(\mathrm{x}),\cdots,u_n(\mathrm{x})$ of $u(\mathrm{x})$, which are defined only on S^{n-1}, to continuous functions $u_i^*(\mathrm{x})$ on D^n, in fact, all of R^n by setting $u^*(0)=0$ and $u^*(\mathrm{x})=ru(\mathrm{x}/r)$ for $r\neq 0$. Let $v(\mathrm{x})=(v_1(\mathrm{x}),\cdots,v_n(\mathrm{x}))$ be a differentiable ϵ-approximation to $u^*(\mathrm{x})$ on D^n with $\epsilon>0$ small enough so that $w(\mathrm{x})=v(\mathrm{x})-(\mathrm{x}\cdot v(\mathrm{x}))\mathrm{x}$ is never zero on S^{n-1}. The vector $w(\mathrm{x})$ for $\mathrm{x}\in S^{n-1}$ is just the orthogonal projection of $v(\mathrm{x})$ onto the tangent hyperplane to S^{n-1} at x; it vanishes only if $v(\mathrm{x})$ is parallel to x, i.e., orthogonal to $u(\mathrm{x})$. Clearly it is possible to choose ϵ so small that this does not occur. Since $w(\mathrm{x})$ is differentiable, tangent to S^{n-1}, and not zero for any $\mathrm{x}\in S^{n-1}$, we have a contradiction to Theorem II′. Thus no vector field $u(\mathrm{x})$ of the type assumed is possible, and Theorem II is established.

5. Related topics. In this section we shall try to give the general reader some insight into the reasons for the importance of Theorems I and II by mentioning a few applications, related theorems, and generalizations. To begin with, we observe that any space X homeomorphic to D^n also has the fixed point property with respect to continuous maps. Indeed, if $H:X\rightarrow D^n$ is a homeomorphism and $F:X\rightarrow X$ is continuous, then $H\circ F\circ H^{-1}:D^n\rightarrow D^n$ has a fixed point x^0; it follows that $y^0=H^{-1}(\mathrm{x}^0)$ is fixed under F. In particular, a compact convex subset of R^n will have the fixed point property; using the Brouwer theorem (Theorem I), it is possible—by a very short and straightforward line of argument—to extend this result to each continuous map of a closed, convex set of a Banach space into a compact subset of itself (Schauder-Mazur theorem). This general-

ized fixed point theorem has been used to prove existence theorems for boundary value problems of differential equations. This will not surprise anyone who has seen proofs of the Inverse Function Theorem and of the Existence Theorem for systems of ordinary differential equations which use the Contracting Mapping Theorem as in [9] and [10]. (This latter theorem is a very simple fixed point theorem which asserts that any continuous map $F: X \to X$ of a complete metric space which contracts distances between pairs of points must have a unique fixed point; although it is not as strong as Theorem I, it is much easier to prove.) A very nice account of the Schauder theorem is found in Bers lecture notes [11], where he gives as an example the following application (C^0 denotes continuous and C^k, $k>0$, k times continuously differentiable functions):

(5.1) *Let $L>0$ be a real number and $f(x_1, x_2, x_3)$ a bounded continuous function on R^3; then there exists a function $x(t)$ of class C^2 which is a solution of the boundary value problem:*

$$x''(t) = f(t, x, x') \quad \textit{and} \quad x(0) = 0 = x(L).$$

To see how this follows from a fixed point theorem one considers the Banach space C_0 of continuous functions $y(t)$ such that $y(0)=0=y(L)$ and its subspaces C_1 and C_2 of functions of class C^1 and C^2 respectively, with suitable norms in all cases. Define $H: C_1 \to C_0$ by $H[x(t)] = f(t, y(t), y'(t))$ and $G: C_0 \to C_2$ by the formula

$$G[y(t)] = \int_0^t \int_0^v y(u)du\, dv - \frac{t}{L} \int_0^L \int_0^v y(u)du\, dv.$$

Then $F = G \circ H$ maps C_1 into $C_2 \subset C_1$ and it is easy to check, using the Fundamental Theorem of Calculus, that $x(t)$ is a fixed point of F if and only if $x(t)$ is the desired solution. Other, similar applications are given in the reference cited [11].

In quite a different direction Theorem I can serve as a cornerstone for dimension theory as may be seen from Brouwer's original work or from the book of that title by Hurewicz and Wallman [12]. As is well known, Brouwer was the first to prove that the dimension of R^n had *topological* meaning by showing that there is no homeomorphism of an open subset of R^n into R^m if $m<n$. In their book Hurewicz and Wallman derive both the fact that R^n is n-dimensional (in a *topologically* invariant sense) and the Lebesgue tiling theorem very easily and directly from the following statement:

(5.2) *Consider $I_n = \{x \in R \mid |x_i| \leqq 1\}$, a cube, and let C_i, C_i' be the opposite faces given by $x_i = +1, -1$ respectively. Let B_i be a closed subset of I_n separating C_i and C_i' for each $i = 1, \cdots, n$. Then there is a point common to all of the B_i.*

We shall reproduce the proof of [12]. Let U_i, U_i' be the components of $I_n - B_i$ and let $f_i(x)$ be the continuous functions on I_n given by $f_i(x) = \pm d(x, B_i)$, the distance of x from B_i with sign depending on whether x is in U_i or U_i'. Then by

Theorem I the map $F: I_n \to I_n$ defined for $\mathbf{x} = (x_1, \cdots, x_n)$ by

$$F(x_1, \cdots, x_n) = (x_1 + f_i(\mathbf{x}), \cdots, x_n + f_n(\mathbf{x}))$$

has a fixed point, which must, then, lie in the intersection of the B_i.

Finally, we mention that Lemma 3, which was the key to proving Theorem I, may be interpreted as stating that any continuous map $F: D^n \to D^n$ which leaves the boundary pointwise fixed must be *onto*. It is possible to relax the condition on the boundary, that is to require that the map F carry the boundary onto itself in an essential way (i.e., not be deformable to a constant map) and still get the same conclusion. This leads directly to the concept of the Brouwer *degree* of a map, an important generalization, which is beautifully developed in [3], using an approach which does not require homology—as does the classical treatment of [8].

Theorem II also has been a well-spring of research in algebraic topology, beginning with the work of Brouwer and Poincaré, then H. Hopf and his students, and continuing to the present day. For example, having answered the question for spheres one might ask exactly which closed differentiable manifolds admit continuous nonvanishing fields of tangent vectors. This question is answered as a special case of the Poincaré-Hopf Theorem, whose full statement requires the definition of the *index* of a vector field at an isolated zero of the field—it is an integer which measures the nature of the singularity. This goes beyond our scope—we refer again to [3]—but we can state the following consequence of the theorem:

(5.3) *A manifold M has a continuous, nonvanishing field of tangent vectors if and only if its Euler characteristic vanishes.*

We will not define the Euler characteristic except to say that in the case of S^{n-1} it has the value 2 for $n-1$ *even* and zero for $n-1$ *odd*—in fact it is zero for any odd dimensional manifold. In the case of surfaces it is zero only for the torus. Therefore the torus is the only closed surface which can have a nonvanishing field of tangent vectors. (Visualizing a closed, orientable surface as a 2-sphere with handles attached, it can be shown that the Euler characteristic is $2(1-g)$, g being the number of handles, see [13, Ch. VI].)

Returning to the case of spheres, a second question which arises naturally is as follows. Given an integer k, $1 \leqq k \leqq n-1$, do there exist k vector fields, $u_1, \cdots, u_k$ tangent to S^{n-1} which not only do not vanish but are even linearly independent at each point? This was a subject of investigation for over thirty years with definitive results being given by J. F. Adams in 1962 [1]. He proved that the *maximum* number of linearly independent vector fields on S^{n-1} was the number which previously had been shown to exist by Hurwitz-Radon-Eckmann; that is, Adams showed that their result was the best possible. The number $\rho(n)-1$ of independent vector fields on S^{n-1} may be defined as follows: write n as an odd integer times a power of 2, i.e., $n=(2a+1)2^b$, and set $b=c+4d$ with $0 \leqq c \leqq 3$. Then the integers a, b, c, d depend only on n, and $\rho(n)$ is given by the

formula $\rho(n)=2^c+8d$. The proof that this many vector fields actually exist depends on purely algebraic results—basically from linear algebra. We shall give a hint of how this very geometric theorem is related to algebra by showing that in the case of S^3 the maximum possible number of independent vector fields can be defined, namely 3. This occurs only in two other cases: S^1 and S^7, as can be checked rather easily by finding all solutions of the equation $\rho(n)-1=n-1$. For the case of S^3, let $\mathbf{x}=(x_1, x_2, x_3, x_4)$ be a point of the unit sphere S^3 in R^4. If we consider it as a vector, then we are required to find three unit vectors $u_i(\mathbf{x})$, $i=1, 2, 3$, which are orthogonal to $\mathbf{x}$, hence tangent to S^3 at $\mathbf{x}$, and which are linearly independent for each $\mathbf{x}$. The following three vectors which are mutually orthogonal clearly satisfy this requirement:

$$u_1 = (-x_2, x_1, -x_4, x_3) \qquad u_2 = (-x_3, x_4, x_1, -x_2) \qquad u_3 = (-x_4, -x_3, x_2, x_1).$$

The interesting fact is that if we identify R^4 with the algebra of quaternions by the correspondence $(x_1, x_2, x_3, x_4)\leftrightarrow x_1+x_2i+x_3j+x_4k$, then these four mutually orthogonal unit vectors are just $\mathbf{x}$, $u_1=i\mathbf{x}$, $u_2=j\mathbf{x}$, and $u_3=k\mathbf{x}$. Using an algebra of dimension 8, the Cayley algebra, precisely the same type of construction gives the 7 vector fields on S^7. Conversely, one has used topological theorems of the type we have been considering to prove the *nonexistence* of certain types of algebras over the real numbers.

Appendix.

PROPOSITION. *For any $\delta>0$ and $\mathbf{x}\in R^n$ there is a C^∞ function $\phi_\delta(y)$ which is non-negative, is zero outside the δ-sphere $\|y\|<\delta$ around the origin and satisfies $\int_{R^n}\phi_\delta(y)dy_1\cdots dy_n=1$.*

Proof. We shall assume that the standard example of C^∞ function given in many calculus texts, e.g., [6, p. 51], is indeed C^∞, and begin from there. The function referred to is

$$\psi(t) = \begin{cases} 0, & t<0 \\ e^{-1/t^2}, & t>0. \end{cases}$$

Then the composite function $\phi_\delta(y)=k\psi(\delta^2-\sum_i y_i^2)$, where k is a positive constant so chosen that

$$\int_{R^n}\phi_\delta(y)dy_1\cdots dy_n = 1,$$

has all the properties claimed.

Proof of the Approximation Lemma for $K=D^n$. Let $f(\mathbf{x})$ be a continuous function on K and extend it to a continuous function on all of R^n by defining it along each radial line outside the boundary sphere to have the same value that it takes on the unit sphere; that is, when $\|\mathbf{x}\|>1$, we define $f(\mathbf{x})$ to be $f(\mathbf{x}/\|\mathbf{x}\|)$. We remark that this is the only place in the proof at which we use the assumption

that K is the unit ball; were we to concede the fact that a continuous function f on an arbitrary compact set $K \subset R^n$ can be extended to a continuous function on all of R^n, then the remainder of the proof would establish the lemma in that case also.

Now, given $\epsilon > 0$, we choose $\delta > 0$ so $x \in K = D^n$ (a compact set) and $\|y\| < \delta$ implies $|f(y+x) - f(x)| < \epsilon$. Define $g(x)$ by

$$g(x) = \int_{R^n} f(y+x)\phi_\delta(y)dy_1 \cdots dy_n.$$

For each fixed x this integral is finite since $\phi_\delta(y) = 0$ if $\|y\| \geqq \delta$, i.e., we could take the region of integration to be the closed δ-ball around x. Moreover,

$$g(x) - f(x) = \int_{R^n} [f(y+x) - f(x)]\phi_\delta(y)dy_1 \cdots dy_n$$

since

$$\int_{R^n} \phi_\delta(y) f(x) dy_1 \cdots dy_n = f(x) \int_{R^n} \phi_\delta(y) dy_1 \cdots dy_n = f(x).$$

Thus $|g(x) - f(x)| < \epsilon \int \phi_\delta(y) dy_1 \cdots dy_n = \epsilon$.

On the other hand, if we change the variable of integration, letting $z = y + x$, we have $g(x) = \int_{R^n} f(z)\phi_\delta(z-x)dz_1 \cdots dz_n$. From this expression, the rule for differentiation under the integral sign, and the fact that ϕ_δ is C^∞, it follows that g is also C^∞.

References

1. J. F. Adams, Vector fields on spheres, Bull. Amer. Math. Soc., 68 (1962) 39–41.
2. R. H. Bing, The elusive fixed point property, this MONTHLY, 76 (1969) 119–131.
3. J. Milnor, Topology from the Differentiable Viewpoint, University of Virginia Press, Charlottesville, 1965.
4. M. Hirsch, A proof of the nonretractability of a cell onto its boundary, Proc. Amer. Math. Soc., 16 (1963) 364–365.
5. H. Flanders, Differential Forms, Academic Press, New York, 1963.
6. W. H. Fleming, Functions of Several Variables, Addison-Wesley, Reading, Mass., 1965.
7. A. Devinatz, Advanced Calculus, Holt, Rinehart & Winston, New York, 1968.
8. Alexandroff-Hopf, Topologie, Springer Verlag, Berlin, 1935.
9. J. Dieudonné, Foundations of Modern Analysis, Academic Press, New York, 1960.
10. Loomis and Sternberg, Advanced Calculus, Addison-Wesley, Reading, Mass., 1968.
11. L. Bers, Topology (Lecture Notes), New York University, 1957.
12. Witold Hurewicz and Henry Wallman, Dimension Theory, Princeton University Press, 1948.
13. David Hilbert and S. Cohn-Vossen, Geometry and the Imagination, Chelsea, New York, 1952.

ON THE PRIME DIVISORS OF POLYNOMIALS

IRVING GERST, State University of New York at Stony Brook, and
JOHN BRILLHART, University of Arizona

1. Introduction. The purpose of this paper is to prove in a fairly elementary way certain results in the theory of higher congruences. These results, which can readily be considered a part of elementary number theory, have previously been established with the use of ideal theory in algebraic number fields. In limiting our means to the basic ideas of field theory and the theory of equations, we hope to make accessible to a fairly wide audience an interesting part of number theory, which is usually not found in beginning treatments of this subject.

Insofar as the theorems presented here are concerned, this paper is purely expository, with the possible exception of the first part of Theorem 11, which we have not seen elsewhere. In other aspects of the paper, however, such as the arrangement of the material, the proofs, and the various applications, we have given a development which we hope will be of interest in itself.

We shall be concerned with those primes p for which the congruence $f(x)\equiv 0 \pmod{p}$ is solvable, where $f(x)$ is a polynomial with rational integer coefficients which is not identically zero (mod p). Such primes will be called *prime divisors* of f, since they divide the value $f(x_0)$ for some integer x_0. To indicate this we write $p\,|\,f(x)$. In this case $f(x)$ also has at least one linear factor (mod p). The set of *all* prime divisors of f will be written $P(f)$.

To illustrate these ideas we recall the following well-known facts: any prime p, $p\nmid a$, will divide $f(x)=ax+b$. Also, for $f(x)=x^2-a$, $P(f)$ can be determined by using the Quadratic Reciprocity Law. The general problem, however, of characterizing the prime divisors of a polynomial of degree >2 is still unsolved, except in certain special cases. (See Hasse [**10**].)

In what follows we present some known results regarding $P(f)$ for various classes of polynomials f. These results are of two types: (1) Infinitely many primes of a certain kind are shown to exist in $P(f)$. For example, it will be shown (Theorem 9) that any nonconstant polynomial possesses infinitely many prime divisors of the form $kn+1$, where n is a given positive integer. (2) Information concerning the sets of primes dividing several polynomials is derived from the relationships that hold between the extensions of the rational field Q defined by these polynomials. For example, if f and g are irreducible over Q with roots α and β respectively, and if $Q(\alpha)=Q(\beta)$, then f and g have the same prime divisors, with at most a finite number of exceptions.

Irving Gerst received his Ph.D. in 1947 at Columbia University under J. F. Ritt in complex variables. From 1945–1961, he was a staff member and then head of applied mathematics groups at Burroughs and RCA. Since 1961 he has been Professor and Chairman of the Department of Applied Analysis at SUNY at Stony Brook, and during 1968 he spent a sabbatical at the University of Arizona. His principal research has been in network theory.

John Brillhart was lecturer and instructor at the University of San Francisco from 1955 to 1965, and an NSF Science Faculty Fellow in 1965–66. He received his Ph.D. at Berkeley under D. H. Lehmer in 1967, and since has been Associate Professor at the University of Arizona. His main areas of research are algebra and number theory. *Editor.*

The usual approach to these results by means of ideal theory is based on the important theorem of Dedekind (cf. [5] or [13, Thm. 8.1, p. 63]) which relates the factorization of $f(x)$ (mod p) to the factorization of p into prime ideals in the algebraic number field defined by a root of $f(x)$. As the ideal theoretic approach is quite powerful and far-reaching in the development of the theory of this subject, we recommend that the reader who is interested in the prime divisors of polynomials also become acquainted with the subject from this point of view. For an introduction to ideal theory in algebraic number fields we suggest Pollard [17]. More recent treatments can be found in Borevich and Shaferevich [4] or Samuel [18]. In the remarks following some of the theorems of the paper, we shall indicate where the reader can find material relevant to the ideal theoretic proofs of these theorems. In referring to this material, which is scattered and rather fragmentary, one must note that the prime divisors of $f(x)$, except for at most a finite number, are exactly those primes which have a prime ideal of the first degree as a factor.

In conclusion, it is important to point out that the methods employed in this paper are constructive and lend themselves to numerical calculation, in that they use the polynomials themselves, their discriminants, and the polynomials relating their roots. This fact allows us to give practical tests to answer various questions concerning polynomials and their fields. For example, it is not difficult to devise a simple and computable sufficient condition for showing that two irreducible polynomials do *not* define the same extensions of Q. (Remark (b) after Theorem 3.)

2. Preliminaries. (A) In this paper we shall always use f, g, and h to denote nonconstant polynomials in one variable with integer coefficients; that is, $f(x) \in Z[x]$. The symbol p will denote a rational prime.

We also use $P_i(f)$, $i=1, 2, \cdots, n$, where $\deg f=n$, to denote the set of prime divisors of f for which the congruence $f(x)\equiv 0 \pmod p$ has *exactly* i distinct solutions. (In this definition a multiple root is counted only once.) Certainly $P(f)$ will then be the union of the sets $P_i(f)$.

A special role is played in this theory by the primes $p\in P(f)$ for which f factors completely into a product of linear factors (mod p). Among such primes are those p for which f will be said to *split completely*.

In defining this special kind of factorization into linear factors we recall first that f (which may be reducible) can be uniquely expressed as $a\prod f_i^{\alpha_i}$, where a, $\alpha_i \in Z$, $\alpha_i>0$, and the $f_i\in Z[x]$ are distinct, primitive, irreducible polynomials with positive leading coefficients [3, pp. 74–76]. Let $g=\prod f_i$. We then say f *splits completely* for $p\in P(f)$ if g is congruent to a product of *distinct* linear factors (mod p), that is, if $p\in P_m(f)$, where $m=\deg g$. (Clearly the maximum number of *distinct* linear factors of f (mod p) is m.) In the important special case where f has no multiple roots, that is, where $n=\deg f=m$, the condition that f splits completely is then $p\in P_n(f)$.

In what follows we shall often be concerned with the comparison of the sets of

primes which divide two or more polynomials. If, for instance, the polynomials f and g have the same prime divisors, except possibly for a finite number, we say they have *essentially* the same prime divisors, and write $P(f)=P(g)$. Thus an $=$ sign, when used in conjunction with "$P(\ )$" or with "$P_i(\ )$", must be understood in this sense. For example, the inequality $P(f)\neq\emptyset$ would mean that $P(f)$ is infinite. We also use $P(f)\subseteq P(g)$ to denote that the prime divisors of f are also divisors of g with the possible exception of a finite number of primes, and say in this case that *almost all* prime divisors of f are divisors of g. (We note that our use here of "$\subseteq$" corresponds to the use of "$\leqq$" by Hasse [**9**, v. 2, p. 141].)

The preceding considerations can be generalized to the case where the coefficients of f and g are integers in an algebraic number field, with p being replaced by a prime ideal in the field. In this elementary introduction, however, we shall restrict our attention solely to rational integral coefficients.

(B) Next recall the relationship between polynomials and fields (cf. [**3**], [**21**]). The fields we shall be dealing with are all finite extensions of the rational field Q (i.e., algebraic number fields). Such extensions will be designated by K, L, and M, and can always be defined (in many ways) by adjoining to Q a root of a polynomial irreducible over Q. (Irreducibility will always be with respect to Q, unless otherwise stated.) Such a root is called a *primitive element* of the extension. We further say a polynomial f *belongs to* K, written $f:K$, if f is irreducible, with a primitive element of K as a root. There exist infinitely many polynomials belonging to any field K.

If $f:K$ and f has the roots $\alpha_1=\alpha, \alpha_2, \cdots, \alpha_n$, where $K=Q(\alpha)$, then $Q(\alpha_i)$ are called the "conjugate fields" of K. If all the conjugate fields are the same, K is said to be *normal* (or Galois). In this case any polynomial belonging to K is also called *normal*. (A normal polynomial is always irreducible.) Evidently when f is normal, the conjugates of a root α can be expressed as polynomials in α with rational coefficients. Conversely, if the roots of an irreducible polynomial f are related to any one of them by such polynomials, then f belongs to the simple, normal extension of Q generated by this root.

Let $K=Q(\alpha)$, $f:K$, $\deg f=n$, and $f(\alpha)=0$. Our proofs will be based, in the main, upon the following two standard results in field theory: (i) *If a polynomial $\phi(x)\in Q[x]$ has α as a root, then $\phi(x)$ is a multiple of $f(x)$ in $Q[x]$.* (ii) *Every element $\gamma\in K$ has a unique representation of the form $\gamma=\sum_{j=1}^{n} r_j\,\alpha^{j-1}$, $r_j\in Q$.*

Additional results in field theory required for our proofs will be introduced later as the need for them arises.

(C) In the sequel we shall always assume a polynomial belonging to a field is *monic*, so that its roots will be algebraic integers. If a given polynomial $f(x)=a_0x^n+\cdots+a_n$, $a_0\neq 0$, is not monic, we can replace it in our considerations by the monic polynomial g, defined by $g(x)=a_0^{n-1}f(x/a_0)$. It is clear from this definition that $P(g)=P(f)$, and also that $P_i(g)=P_i(f)$ for $i=1, 2, \cdots, n$, since the only difference between the two sets of primes is at most the finite number of primes dividing a_0. It is also clear that if $f:K$, then $g:K$. These

observations imply that all theorems proved for monic polynomials in this paper are also true for nonmonic polynomials.

We conclude this section by recalling several elementary matters in the arithmetic of rational numbers (mod p). A fraction a/b (mod p) is defined for $p\nmid b$ as the unique solution x (mod p) of the congruence $bx\equiv a$ (mod p). The usual laws for operating with congruences continue to hold for these fractions. Moreover, if $p\nmid b$ and $f(a/b)=A/B$, A, $B\in Z$, and $p\nmid B$, then $f(a/b)\equiv A/B$ (mod p), where a/b and A/B are interpreted as fractions (mod p). In particular, if $p|A$ then $f(a/b)\equiv 0$ (mod p) and $p\in P(f)$. This particular argument will appear in many of the proofs.

3. Basic theorems. In this section we give three theorems which will be fundamental in much that follows. The first theorem is quite well known.

THEOREM 1 (Schur [20]). *Every nonconstant polynomial f has an infinite number of prime divisors; that is, $P(f)\neq\varnothing$.*

Proof. We may suppose $f(0)=c$ is not zero, since otherwise every prime divides $f(x)$. Furthermore, there exists at least one prime divisor of $f(x)$, for a polynomial can take the values ± 1 at most a finite number of times. Assume $p_1, p_2, \cdots, p_n$ are the only prime divisors of $f(x)$ and let $a=p_1p_2\cdots p_n$. Then $f(acx)=cg(x)$, where $g(x)=1+c_1x+c_2x^2+\cdots$, and $a|c_i$. A prime divisor p of $g(x)$ will also divide $f(x)$, and so must be one of the p_i. But then $p|c_i$, which implies $p|1$. Hence, $g(x)$ has no prime divisors, i.e., $g(x)=\pm 1$ for all integral x, which is impossible. ∎

In the second theorem we show that the primes dividing a polynomial belonging to a field essentially contain those that divide a polynomial belonging to any extension of that field.

THEOREM 2. *If $K\subseteq L$, $f{:}K$, and $g{:}L$, then $P(f)\supseteq P(g)$.*

Proof. We must show that almost all the prime divisors of g are also divisors of f. Let $f(\alpha)=g(\beta)=0$, $\alpha\in K$, $\beta\in L$. Then, since $K\subseteq L$, we can write $\alpha=\phi(\beta)$, $\phi(x)\in Q[x]$. But then $f(\phi(x))$ and $g(x)$ have β as a common root, so the irreducibility of $g(x)$ implies

$$f(\phi(x)) = g(x)g_1(x),\ g_1(x)\in Q[x]. \tag{1}$$

Now suppose $p\in P(g)$ does not divide any of the denominators in the coefficients of ϕ (assumed to be in lowest terms). (Since $g(x)$ is monic by assumption (cf. Sec. 2), the same will be true for the coefficients of $g_1(x)$.) Then from (1) if $g(b)\equiv 0$ (mod p), $b\in Z$, $\phi(b)$ will be a root of $f(x)\equiv 0$ (mod p). Thus $p|f(x)$. ∎

REMARKS. (a) It can be shown using ideal theory (Schinzel [19]), that there exist K and L, K a *proper* subfield of L, for which nonetheless $P(f)=P(g)$.

(b) For a discussion using ideal theory relating to Theorem 2, see [12, Prop. 20, p. 19].

Theorem 2 may be of use in showing a field K defined by f is not contained in

a field L defined by g, $\deg f \leqq \deg g$. (Certainly $K \not\subseteq L$ if $\deg f \nmid \deg g$.) For, if $g(\beta)=0$, consider all root polynomials $\phi(x)$, where $\phi(\beta)$ may represent some root α of f. If a prime p can be found such that (i) $p \mid g(x)$, (ii) $p \notin P(f)$, and (iii) p does not divide any denominator in all of the $\phi(x)$, then $K \not\subseteq L$.

From a computational point of view it is clear how one would attempt to find a prime p satisfying (i) and (ii). (See example following Theorem 3.)

To verify whether such a p also satisfies (iii), we need to identify the primes which may possibly occur in the denominators of $\phi(x)$. This is accomplished in the following Lemma.

LEMMA 1. *Let $L=Q(\beta)$ where β is an algebraic integer of degree n over Q with defining polynomial g. Also, let the canonical field representation of any algebraic integer $\alpha \in L$ be $\alpha=\phi(\beta)=\sum_{j=1}^{n} r_j\beta^{j-1}$ where the $r_j \in Q$ are in lowest terms. If p divides the denominator of any r_j, then it must divide the discriminant $D(g)$ of g.*

Proof. In establishing this lemma, which will be used in numerical applications, we shall require somewhat more than the rudiments of field theory, in that we assume the basis theorem for the domain of integers in an algebraic number field (see [**17**, p. 65]):

There exist algebraic integers $\gamma_i \in L$, $i=1, 2, \cdots, n$, such that if $\alpha \in L$ is an algebraic integer, then α has a unique representation of the form

$$\alpha = \sum_{i=1}^{n} d_i\gamma_i, \qquad d_i \in Z. \tag{2}$$

Since the powers of β are also integers in L,

$$\beta^{j-1} = \sum_{i=1}^{n} c_{ji}\gamma_i, \quad c_{ji} \in Z, \qquad j = 1, 2, \cdots, n. \tag{3}$$

Denote the field conjugates of β and the γ_i respectively as $\beta=\beta_1, \beta_2, \cdots, \beta_n$ and $\gamma_i=\gamma_i^{(1)}, \gamma_i^{(2)}, \cdots, \gamma_i^{(n)}$, $i=1, 2, \cdots, n$.

Using (3) we can obtain the *matrix* relation

$$(\beta_j^{i-1}) = (c_{ij})(\gamma_i^{(j)}), \qquad i, j = 1, 2, \cdots, n.$$

Taking the determinant of each side and squaring we obtain on the left the square of the Vandermonde determinant, which is well known to be the discriminant D of g. Thus,

$$D(g) = J^2 \cdot \Delta, \tag{4}$$

where $J=\det(c_{ij})$ and $\Delta=[\det(\gamma_i^{(j)})]^2$. In this equation J and $D(g)\neq 0$ are rational integers, and since Δ is thus an algebraic integer equal to a rational number, it must also be a rational integer.

It is worth mentioning that $|J|$ and Δ, customarily called the "index" of β and the "discriminant" of L respectively, can be shown to be independent of the choice of the basis $\{\gamma_i\}$.

Next, using (3) in $\alpha = \sum_{j=1}^{n} r_j \beta^{j-1}$, we get

$$\alpha = \sum_{i=1}^{n} \left(\sum_{j=1}^{n} r_j c_{ji} \right) \gamma_i.$$

Comparing this equation with (2) and equating corresponding coefficients of γ_i yields the system

$$d_i = \sum_{j=1}^{n} r_j c_{ji}, \qquad i = 1, 2, \cdots, n.$$

Solving for r_j (which is in lowest terms), $r_j = S_j/J$, $j = 1, 2, \cdots, n$, where the rational integers S_j are the numerators in Cramer's rule. Thus, any prime p dividing the denominator of r_j must divide J, which by (4) gives $p \mid D(g)$. ∎

REMARKS. (a) We observe from the above proof that Lemma 1 could have been stated with the index $|J|$ replacing $D(g)$. However, since $D(g)$ is much easier to compute (mod p) than the index, it is sufficient in numerical applications to use the lemma as stated.

(b) In ideal theoretic terms the condition $p \nmid D(g)$ implies that p is "unramified" in L, that is, p factors into distinct prime ideals in L.

The final theorem of this section gives information about the sets of primes that divide two different polynomials belonging to the *same* field.

THEOREM 3. *If $f:K$ and $g:K$, then $P(f) = P(g)$ and $P_i(f) = P_i(g)$ for $i = 1, 2, \cdots, n$, where $\deg f = \deg g = n$.*

Proof. By Theorem 2, $P(f) \subseteq P(g)$ and $P(f) \supseteq P(g)$. Hence $P(f) = P(g)$.

To prove the second part of the theorem, we first construct a 1-1 mapping of the incongruent roots of $g(x) \equiv 0 \pmod p$ *into* the set of incongruent roots of $f(x) \equiv 0 \pmod p$ for almost all primes p dividing both f and g.

Suppose $f(\alpha) = g(\beta) = 0$, $\alpha, \beta \in K$. Then there exist polynomials $\phi(x)$, $\psi(x) \in Q[x]$ for which $\beta = \psi(\alpha)$ and $\alpha = \phi(\beta)$. With this ϕ, define the map: $b_i \to \phi(b_i)$, where the b_i are the incongruent roots of $g(x) \equiv 0 \pmod p$. As in the proof of Theorem 2, we can derive both identity (1), and, if p does not divide the denominators in ϕ, the conclusion that follows from (1), which shows $\phi(b_i)$ is also a root of $f(x) \equiv 0 \pmod p$.

Next we show that $\phi(b_i) \not\equiv \phi(b_j) \pmod p$, $i \neq j$. From the equation $\psi(\phi(\beta)) = \beta$ we obtain the identity

$$\psi(\phi(x)) - x = g(x) g_2(x), \qquad g_2(x) \in Q[x]. \tag{5}$$

Suppose that $\phi(b_i) \equiv \phi(b_j) \pmod p$. Then setting $x = b_i$ and b_j in (5), if p does not divide the denominators in ψ, we find $b_i \equiv \psi(\phi(b_i)) \equiv \psi(\phi(b_j)) \equiv b_j \pmod p$, which is a contradiction. If we now reverse the roles of f and g in the preceding argument, we obtain a 1-1 mapping between the two sets of incongruent roots. ∎

REMARKS. (a) Since Theorem 3 is implied by Dedekind's theorem, ideal

theoretic proofs can be found in the references cited in the introduction.

(b) It is clear from Lemma 1 and the proof of Theorem 3, if $p \nmid D(f)D(g)$, that $p|f(x)$ if and only if $p|g(x)$. Thus, by way of application, to show two irreducible polynomials of the same degree do *not* determine the same field, we have only to find a prime p such that $p \nmid D(f)D(g)$, $p|f(x)$, and either $p \nmid g(x)$, or, if $p|g(x)$, then f and g do not have the same number of incongruent roots (mod p).

Example. Consider the two irreducible polynomials $f_1(x) = x^4 + 4x^3 - 4x^2 - 40x - 56$ and $f_2(x) = x^4 - 8x^2 - 24x - 20$, both with discriminant $D = -2^{12} \cdot 3^2 \cdot 31$. (Irreducibility can be shown by using undetermined coefficients and simple parity arguments.) We note $13|f_1(2) = -2^3 \cdot 13$, but $13 \nmid f_2(x)$ for $x = 0, 1, \cdots, 12$. Also $13 \nmid D$. Hence, the polynomials belong to different fields.

4. Polynomials belonging to normal fields. We next investigate the behavior of a *normal* polynomial with respect to its prime divisors.

THEOREM 4. *A normal polynomial splits completely for almost all its prime divisors.*

Proof. Let $f(x) = x^n + a_1x^{n-1} + \cdots + a_n$ be normal. Since $P_n(f) \subseteq P(f)$, we have only to show $P(f) \subseteq P_n(f)$. From the normality of f its roots can be written as

$$\alpha_1 = \alpha, \alpha_2 = \phi_2(\alpha), \cdots, \alpha_n = \phi_n(\alpha), \qquad \phi_i(x) \in Q[x].$$

Using these ϕ_i a useful two-variable identity can be obtained by expanding the product

$$(6) \quad (y - x)(y - \phi_2(x)) \cdots (y - \phi_n(x)) = y^n + r_1(x)y^{n-1} + \cdots + r_n(x).$$

If we set $x = \alpha$ in (6), the left side becomes $f(y)$, while the coefficients on the right reduce to a_i. Hence, $r_i(x) - a_i$ and $f(x)$ have α as a common root, from which the irreducibility of f implies $r_i(x) - a_i = f(x)g_i(x)$, $i = 1, 2, \cdots, n$, $g_i(x) \in Q[x]$. Replacing the $r_i(x)$ in (6) with these results, we obtain the desired identity

$$(7) \; (y - x)(y - \phi_2(x)) \cdots (y - \phi_n(x)) = f(y) + f(x)g_1(x, y), \; g_1(x, y) \in Q[x, y].$$

Now suppose p does not divide any of the denominators in the ϕ_i (and hence also in g_1). Also, let $p \in P(f)$; say $p|f(a)$. Then, substituting $x = a$ in (7) and considering (7) as a congruence (mod p), we have

$$(8) \qquad (y - a)(y - \phi_2(a)) \cdots (y - \phi_n(a)) \equiv f(y) \pmod{p}.$$

If $p \nmid D(f)$ as well, the roots on the left side of (8) are distinct, which implies $p \in P_n(f)$. ∎

REMARKS. (a) For an ideal theoretic proof of Theorem 4 see Lang [**12**, Cor. 2, p. 21] or Mann [**13**, p. 69].

(b) Note by virtue of Lemma 1 the single condition $p \nmid D(f)$ encompasses

all the restrictions on p in the above proof, and consequently f will split completely for such $p \in P(f)$.

By way of application, to show that a given irreducible polynomial f is *not* normal, we need only find a prime $p \in P(f)$, $p \nmid D(f)$, for which f does not split completely.

Example. Let $f(x)=x^4-2x^3-2x^2+3x+1$. Then $f(x)$ is irreducible, since it is irreducible (mod 2). Now $f(5)=11\cdot 31$ and $11 \nmid D(f)=5^2\cdot 29$. With $p=11$, $f(x) \equiv (x-5)(x-7)(x^2-x-5)$ (mod 11), where the quadratic factor is irreducible (mod 11). Thus f is not normal.

(c) It is worth mentioning in the above proof that the functions ϕ_i, which relate the roots of f over the rationals, are the same functions reduced (mod p), which relate the roots of f in Z_p, the integers (mod p).

Example. The polynomial $f(x)=x^3-3x+1$ is normal with roots α, $\phi_2(\alpha) =\alpha^2-2$, and $\phi_3(\alpha)=-\alpha^2-\alpha+2$, where α is any root of $f(x)$. Also, $D(f)=3^4$. Now $f(8)=3\cdot 163$, so we can choose $p=163$ since $163 \nmid D(f)$. The two remaining roots of $f(x)$ (mod 163) can then be obtained by putting $a=8$ in ϕ_2 and ϕ_3, giving the complete factorization

$$x^3 - 3x + 1 \equiv (x - 8)(x - 62)(x - 93) (\text{mod } 163).$$

In practice, if it happens that $f(x)$ splits completely for each prime divisor tried, we have a good indication that f is normal, since the converse of Theorem 4 is true [**9**, vol. 2, p. 141].

(d) Theorem 4 is a special case of the following more general theorem, which shows how a normal polynomial f factors (mod p), where p does not necessarily divide $f(x)$.

THEOREM. *If f is a normal polynomial and $p \nmid D(f)$, then f factors (mod p) into irreducible factors of the same degree, where the degree depends on p.*

An elementary proof of this result can be given using (7). However, we omit the proof here, because of the emphasis of this paper on linear factors. We note, however, the following interesting result:

COROLLARY. *If f is a normal polynomial of prime degree and $p \nmid D(f)$, then f is either irreducible (mod p) or splits completely (mod p).*

We next consider the special case of Theorem 2 in which the overfield L is assumed to be normal. In this case we can obtain more precise information about the prime divisors of g.

THEOREM 5. *If $K \subseteq L$ with L normal, $f{:}K$, and $g{:}L$, then $P_n(f) \supseteq P(g)$, where deg $f=n$.*

Proof. We must show that f splits completely for almost all the prime divisors of g. By hypothesis the roots of f can be written as $\alpha_1=\phi_1(\beta)$, $\alpha_2=\phi_2(\beta)$, $\cdots$,

$\alpha_n=\phi_n(\beta)$, where β is any root of g and $\phi_i(x)\in Q[x]$. Using the same arguments as in the proof of Theorem 4, we can establish the identity

$$(9)\qquad \prod_{i=1}^{n}[y-\phi_i(x)]=f(y)+g(x)g_1(x,y),\qquad g_1(x,y)\in Q[x,y].$$

As before, if $g(b)\equiv 0 \pmod{p}$, $b\in Z$, but p does not divide any denominator in the ϕ_i, then

$$f(y)\equiv\prod_{i=1}^{n}[y-\phi_i(b)]\qquad(\operatorname{mod} p),$$

where if $p\nmid D(f)$, the roots on the right side are distinct (mod p). ∎

COROLLARY. *Every nonconstant polynomial has an infinite number of prime divisors for which it splits completely.*

Proof. Consider the normal extension $L=Q(\alpha_1,\cdots,\alpha_n)$, where α_i are the roots of the given polynomial f. Let $g:L$. Each irreducible factor f_i of f over Z (as defined in Section 2(A)) defines a simple extension of Q which is a subfield of L. From Theorem 5 applied to these subfields, each f_i will split completely for almost all prime divisors p of g. By Theorem 1, g possesses infinitely many prime divisors. Hence, omitting from $P(g)$ the finite number of p which divide $D(h)(\neq 0)$, where $h(x)$ is the product of all the f_i, the corollary follows. ∎

We now inquire whether the inclusion in Theorem 5 can be reversed. Although this cannot be done in general for *any* normal extension, the reversal is legitimate in the following special case.

THEOREM 6. *Let $K\subseteq L$ with L the smallest normal extension of K. If $f:K$ and $g:L$, then $P_n(f)=P(g)$, where $\deg f=n$.*

REMARK. The field L can be described, alternatively, as the splitting field of f.

Proof. By Theorem 5 we need only prove $P_n(f)\subseteq P(g)$. Suppose $\alpha_1,\cdots,\alpha_n$ are the roots of f and $p\in P_n(f)$. By hypothesis we then have

$$(10)\qquad f(x)=\prod_{i=1}^{n}(x-\alpha_i)\equiv\prod_{i=1}^{n}(x-a_i)\quad(\operatorname{mod} p),$$

with the $a_i\in Z$ distinct. From (10) we conclude

$$(11)\qquad G_j(\alpha_i)\equiv G_j(a_i)\quad(\operatorname{mod} p),\qquad j=1,2,\cdots,n,$$

where G_j denotes the jth elementary symmetric function.

Next consider the construction of L from K. It is well known for the *smallest* normal extension of K that there exists a primitive element $\beta\in L$ of the form $\beta=\sum_{i=1}^{n}c_i\alpha_i$, where the $c_i\in Z$. Without loss of generality we can suppose that β is already a root of g, since by Theorem 3, any two polynomials belonging to the same field have essentially the same prime divisors. Now form the two polynomials

$$h(x) = \prod_{s_j \in S_n} \left[x - s_j\left(\sum_i c_i\alpha_i \right) \right] = \sum_k d_k x^k, \quad \text{where} \quad d_{n!} = 1$$

and

$$h^*(x) = \prod_{s_j \in S_n} \left[x - s_j\left(\sum_i c_i a_i \right) \right] = \sum_k d_k^* x^k, \quad \text{where} \quad d_{n!}^* = 1.$$

In these two products the s_j range over all permutations of the symmetric group S_n on n letters, where for example, by $s_j(\sum_i c_i\alpha_i)$ we mean $\sum_i c_i\alpha_l$, $l = s_j(i)$.

The coefficients d_k, as symmetric functions of the α_i, are of course integers, and may be written as $d_k = F_k(G_1(\alpha_i), G_2(\alpha_i), \cdots)$, where F_k is a polynomial with integer coefficients. By construction the coefficients d_k^* of $h^*(x)$ are determined by exactly the same polynomials F_k, with the arguments $G_j(\alpha_i)$ replaced by $G_j(a_i)$. It then follows from (11) that $d_k^* \equiv d_k \pmod{p}$, and therefore

$$h(x) \equiv h^*(x) \pmod{p}. \tag{12}$$

Since $h^*(x)$ is a product of linear factors (mod p), the same is true for $h(x)$ (mod p). But $g(x)$ has the root β in common with $h(x)$. Hence, $g(x)$ divides $h(x)$, so $p \mid g(x)$. ∎

REMARKS. (a) For material relating to an ideal theoretic handling of Theorem 6 see Mann [**13**, Cor. 13.5.1, p. 113].

(b) From the last sentence of the preceding proof (possibly renumbering the s_j) we get that

$$g(x) \equiv \prod_{j=1}^{m} \left[x - s_j\left(\sum_i c_i a_i \right) \right] \pmod{p}, \qquad m = \deg g.$$

In this product it can be shown that s_j ranges over those permutations of S_n for which $\beta_j = s_j(\sum_i c_i\alpha_i)$ are the m conjugates of β. (These s_j, of course, constitute the Galois group of f.) The roots of $g(x)$ (mod p) are then given in terms of the roots of $f(x)$ (mod p) in the same way the roots of $g(x)$ are given in terms of the roots of $f(x)$ over Q.

Example. Let $f(x) = x^3 - 2$ with roots $\alpha_1 = \theta$, $\alpha_2 = \omega\theta$, and $\alpha_3 = \omega^2\theta$, $\theta = 2^{1/3}$ and $\omega = [-1 + (-3)^{1/2}]/2$. The smallest normal extension of Q containing the α_i is of degree 6, since f is not normal. (Remark (b) following Theorem 4 with $p = 5$ and $x = 3$.) The expression $\zeta = \alpha_1 + a\alpha_2 + b\alpha_3 = \theta(1 + a\omega + b\omega^2)$, a, $b \in Z$ is now chosen so that ζ will take 6 different values under the 6 permutations of the α_i. For example, let $a = 2$ and $b = -1$. Then $\zeta = \theta(1 + 2\omega - \omega^2)$ is a root of the (resolvent) equation $g(x) = x^6 + 40x^3 + 1372 = 0$. Now $31 \in P_3(f)$ since $f(x) \equiv (x-4) \cdot (x-7)(x-20) \pmod{31}$. The 6 roots b_i of $g(x) \equiv 0 \pmod{31}$ can now be computed by carrying out the 6 permutations on $a_1 \equiv 4$, $a_2 \equiv 7$, and $a_3 \equiv 20 \pmod{31}$ in the expression $a_1 + 2a_2 - a_3$; that is, $(a_1, a_2, a_3) = (4, 7, 20)$ gives $b_1 \equiv -2$, (a_1, a_3, a_2) gives $b_2 \equiv 6 \pmod{31}$, etc. Finally,

$$g(x) \equiv (x-6)(x-12)(x-21)(x-26)(x-29)(x-30) \pmod{31}.$$

(c) In Theorem 6 no information about $P_i(f)$ is obtainable for $i=1, 2, \cdots, n-2$, since all but at most a finite number of primes in these sets lie outside of $P(g)$.

(d) If in Theorem 6 the field K is already normal, then $K=L$ and Theorem 6 reduces to Theorem 4.

5. Applications. We are now in a position to prove several striking results about the prime divisors of arbitrary polynomials. As far as we know, these theorems, due to Nagell [15], have been proved previously only with the use of ideal theory. (See also Fjellstedt [6].)

THEOREM 7. *For any two nonconstant polynomials f and g, there exist infinitely many primes which split both f and g completely.*

Proof. Adjoin all the roots of f and g to Q, forming the normal field L. Let $h:L$. Then each irreducible factor of f or g over Z defines a simple extension of Q, which is a subfield of L. As in the proof of the Corollary to Theorem 5, it now follows that f and g individually split completely for almost all of the prime divisors of h. But then there must exist infinitely many primes for which, simultaneously, f and g split completely. ∎

COROLLARY. *Theorem 7 holds for a finite number of nonconstant polynomials.*

REMARK. Recently Nagell [16], by an elementary method different from our own, has proved the weaker result that there exist infinitely many common prime divisors of any finite number of nonconstant polynomials.

We next quote the well-known theorem on the form of the prime divisors of the cyclotomic polynomial $Q_n(x)$, that is, the monic polynomial whose roots are the primitive nth roots of unity [14, Th. 94, p. 164].

THEOREM 8. *If $p \nmid n$, then $p \mid Q_n(x)$ if and only if $p \equiv 1 \pmod{n}$.*

As an application, if in the Corollary to Theorem 7 we consider the set of $m+1$ polynomials $f_1, \cdots, f_m, Q_n$ and use Theorem 8, we obtain Theorem 9.

THEOREM 9. *If $f_1, \cdots, f_m$, $m \geqq 1$, are any nonconstant polynomials, then for each fixed $n \geqq 1$ there exist infinitely many primes of the form $kn+1$ for which each of the f_i splits completely.*

6. Polynomials belonging to composite fields. In this section we consider a more complicated situation than previously, in that we study the prime divisors of polynomials belonging to composite fields.

Recall the following facts about such fields: If $K=Q(\alpha)$ and $L=Q(\beta)$ are two finite, simple extensions of Q, then the composite field $M=KL$ is defined as $Q(\alpha, \beta)$. The field M is also a finite, simple extension of Q, for which a primitive element γ exists in the form $\gamma=\alpha+a\beta$, where a is a suitably chosen rational integer. The defining polynomial $h(y)$ of γ will then be the irreducible factor of

$$v(y) = \prod_{i=1}^{n} \prod_{j=1}^{m} [y - (\alpha_i + a\beta_j)], \qquad v(y) \in Q[y], \tag{13}$$

which has γ as a root. In this definition the α_i and β_j are the conjugates of $\alpha_1 = \alpha$ and $\beta_1 = \beta$, and a is chosen so that $v(y)$ has no multiple roots.

In particular, it follows if K and L are normal that $M = KL$ will also be normal. To see this, observe that any conjugate γ_k of γ will be of the form $\gamma_k = \alpha_i + a\beta_j$ for some i and j. The normality of K and L then implies the particular roots α_i and β_j must lie in K and L respectively. But K and L are subfields of M, so α_i and β_j, and hence γ_k are rationally expressible in terms of γ. Hence, M is normal.

THEOREM 10. *Let K and L be normal and let $M = KL$. If $f:K$, $g:L$, and $h:M$, then $P(h) = P(f) \cap P(g)$.*

Proof. Since $K \subseteq M$ and $L \subseteq M$, it follows from Theorem 2 that $P(f) \supseteq P(h)$ and $P(g) \supseteq P(h)$. Thus, $P(f) \cap P(g) \supseteq P(h)$. It remains to show that $P(f) \cap P(g) \subseteq P(h)$. If $\alpha_1 = \alpha, \alpha_2, \cdots, \alpha_n$ and $\beta_1 = \beta, \beta_2, \cdots, \beta_m$ are the roots of f and g respectively, then the normality of f and g implies there exist polynomials $\phi_i(x), \psi_j(x) \in Q[x]$ such that $\alpha_i = \phi_i(\alpha)$, $i = 1, 2, \cdots, n$, and $\beta_j = \psi_j(\beta)$, $j = 1, 2, \cdots, m$, where we take $\phi_1(x) = \psi_1(x) = x$. As previously, assume $\gamma = \alpha + a\beta$ is already a root of h, a a suitably chosen integer.

We next establish an identity analogous to (9), involving both f and g. If in identity (7) y is replaced by $y - a\psi_j(z)$, z a new indeterminate, we obtain

$$\prod_{i=1}^{n} [y - a\psi_j(z) - \phi_i(x)] = f(y - a\psi_j(z)) + f(x)q_j(x, y, z),$$

$q_j(x, y, z) \in Q[x, y, z]$. If j ranges over the values $1, 2, \cdots, m$ and the resulting identities are multiplied together, then

$$\prod_{j=1}^{m} \prod_{i=1}^{n} [y - a\psi_j(z) - \phi_i(x)] = \prod_{j=1}^{m} f(y - a\psi_j(z)) + f(x)r(x, y, z), \tag{14}$$

$r(x, y, z) \in Q[x, y, z]$. Expanding the product on the right gives

$$\prod_{j=1}^{m} f(y - a\psi_j(z)) = y^{mn} + r_1(z)y^{mn-1} + \cdots + r_{mn}(z),$$

$r_i(z) \in Q[z]$. Since substituting $z = \beta$ in the left side of this expansion gives

$$\prod_{j=1}^{m} f(y - a\beta_j) = \prod_{j=1}^{m} \prod_{i=1}^{n} [(y - a\beta_j) - \alpha_i] = v(y)$$

by (13), we can proceed as we did in establishing (7) to obtain

$$\prod_{j=1}^{m} f(y - a\psi_j(z)) = v(y) + g(z)s(y, z), \qquad s(y, z) \in Q[y, z].$$

Putting this result in (14), we obtain the desired identity

$$\prod_{j=1}^{m}\prod_{i=1}^{n}[y-\phi_i(x)-a\psi_j(z)] = v(y)+f(x)r(x,y,z)+g(z)s(y,z). \tag{15}$$

Now suppose $p\in P(f)\cap P(g)$ and p does not divide any denominator in (15); that is, there exist b, $c\in Z$ for which $f(b)\equiv g(c)\equiv 0 \pmod p$. Then it is clear $v(y)$ splits into (not necessarily distinct) linear factors (mod p) when we put $x=b$ and $z=c$ in (15), and consider (15) (mod p). Since $h(y)$ is a factor of $v(y)$, it follows that $p\in P(h)$. ∎

REMARKS. (a) For an ideal theoretic proof see Hasse [**9**, vol. 1, p. 50].

(b) It can be readily established by examining the above proof that the only primes which need to be excluded as divisors of the denominators in (15) are those dividing $D(f)D(g)$.

Example. Let $f(x)=x^2+1$ and $g(x)=x^2-2$. The corresponding (normal) fields are $K=Q(i)$ and $L=Q(2^{1/2})$. If, say, we choose $\gamma=2^{1/2}+i$ as a primitive element for KL, then the defining polynomial for γ is $h(x)=x^4-2x^2+9$. We shall determine $P(h)$.

We have by the first and second supplements to the Quadratic Reciprocity Law that

$$P(f) = \{p \mid p = 2 \text{ or } p\equiv 1 \pmod 4\}$$

and

$$P(g) = \{p \mid p = 2 \text{ or } p\equiv \pm 1 \pmod 8\}.$$

As $D(f)D(g)=-32$, it follows from Remark (b) that every prime in $P(f)\cap P(g)$ with the possible exception of $p=2$ is in $P(h)$. By inspection, we see that also $2\in P(h)$. (Observe the prime 2, though it is a divisor of both $D(f)$ and $D(g)$, still occurs as a divisor of f, g, and h.) On the other hand, since $D(h)=2^{14}\cdot 3^2$, $p=3$ is the only possible prime divisor of $h(x)$ which is not also in $P(f)\cap P(g)$. Evidently $3\mid h(x)$ and so

$$P(h) = \{p \mid p = 2,\ 3 \text{ or } p\equiv 1 \pmod 8\}.$$

That this result is correct can also be seen from Theorems 3 and 8 and the observation that $Q(\gamma)=Q(\gamma')$ where $\gamma'=(1+i)/2^{1/2}$ is a root of the cyclotomic polynomial $Q_8(x)=x^4+1$. Here $D(Q_8)=2^8$, which implies that $P(h)$ and $P(Q_8)$ are the same except for the possible existence of prime divisors $p=2$ or $p=3$ which may be present in just one of these sets.

COROLLARY. *If $f:K$, $g:L$, and $h:M$ with $\deg f=n$, $\deg g=m$, and $\deg h=k$, where $M=KL$, then $P_k(h)=P_n(f)\cap P_m(g)$.*

Proof. Let K_1, L_1, and M_1 be the *smallest* normal extensions of K, L, and M respectively, and let $f_1:K_1$, $g_1:L_1$, and $h_1:M_1$. Then by Theorem 6, $P_n(f)=P(f_1)$,

$P_m(g)=P(g_1)$, and $P_k(h)=P(h_1)$. Since it can be shown that $M_1=K_1L_1$, the Corollary follows from Theorem 10. ∎

We now present a generalization of Theorem 10, in which only one of the fields K and L is normal. We have presented Theorem 10 separately, however, since the proof explicitly exhibits the solutions of the congruences involved.

THEOREM 11. *Let $M=KL$, where K is normal, and let $f:K$, $g:L$, and $h:M$, where $\deg g=m$. Also, if $\alpha\in K$ is a root of f, let the degree of α relative to L be s. Then $\deg h=ms$, and for $1\leqq i\leqq ms$ we have $P_i(h)=\emptyset$, if $s\nmid i$, and $P_i(h)=P(f)\cap P_r(g)$, if $i=rs$, $1\leqq r\leqq m$. Also, $P(h)=P(f)\cap P(g)$.*

Proof. Let $\alpha_1=\alpha, \alpha_2, \cdots, \alpha_n$ and $\beta_1=\beta, \beta_2, \cdots, \beta_m$ be the roots of f and g respectively. Also, let $\alpha+a\beta$, $a\in Z$, be a root of h, and $L=Q(\beta)$. Consider the factorization

$$f(x)=\prod_{k=1}^{t} f_k(x;\beta), \tag{16}$$

where the f_k are monic and irreducible over L. If f_1 has α as a root, then by hypothesis $\deg f_1=s$. Thus, since $M=L(\alpha)$, it follows that $\deg h=ms$. The normality of f also implies $M=L(\alpha_i)$ for each i, so the degree of all the α_i relative to L is s. Hence, the degree of all the f_k is s, and thus, comparing the degrees of the two sides of (16), we have $n=st$.

A factorization like that in (16) holds if β is replaced by any of its conjugates, since the fields $Q(\beta_j)$ are isomorphic. Thus, substituting β_j for β and $x-a\beta_j$ for x in (16), we obtain

$$f(x-a\beta_j)=\prod_{k=1}^{t} f_k(x-a\beta_j;\beta_j), \qquad j=1,2,\cdots,m.$$

Multiplying these m equations together and using (13) gives

$$v(x)=\prod_{j=1}^{m} f(x-a\beta_j)=\prod_{k=1}^{t} h_k(x),$$

where $h_k(x)=\prod_{j=1}^{m} f_k(x-a\beta_j;\beta_j)$. Since this latter product is symmetric in the β_j, $h_k(x)\in Z[x]$, where $\deg h_k=ms$. Clearly $h_1(x)=h(x)$. Thus, all the $h_k(x)$ belong to M, for certainly $\alpha_i+a\beta$ will be a root of $h_k(x)$ for some i, so, by the normality of f, $Q(\alpha_i+a\beta)=Q(\alpha+a\beta)=M$. Consequently, the $h_k(x)$ are irreducible over Q for all k.

Next consider the respective roots α_i^*, $i=1,2,\cdots,n$, and β_j^*, $j=1,2,\cdots,m$ of f and g over Z_p. Forming the polynomial

$$v^*(x)=\prod_{i=1}^{n}\prod_{j=1}^{m}(x-\alpha_i^*-a\beta_j^*),$$

we have $v^*(x)\in Z_p[x]$, and then, by the argument used to establish (12), we

find

$$v(x) \equiv v^*(x) \pmod{p}. \tag{17}$$

Also, if $p \nmid D(f)D(g)$ and $p \in P(f) \cap P_r(g)$ for some r, $1 \leqq r \leqq m$, then exactly r distinct β_j^* will lie in Z_p, and by Theorem 4, all the α_i^* will be distinct in Z_p. Hence, if $p \nmid D(v)$, v^* will have at least nr distinct linear factors (mod p). But v^* cannot have more than this number, for the presence of a linear factor implies for some i and j that $\alpha_i^* + a\beta_j^* \in Z_p$, and since the $\alpha_i^* \in Z_p$ and $p \nmid a$, then $\beta_j^* \in Z_p$. (Actually $p \nmid a$ follows directly from $p \nmid D(v)$.)

Now by (17), $p \mid v(x)$, and hence p divides at least one of the factors $h_k(x)$ of $v(x)$. Say $p \in P_i(h_k)$. Then by Theorem 3, almost all the primes dividing this factor will also divide every other factor of $v(x)$, and the number of incongruent solutions (mod p) in each of these will be the same. Hence, the number of distinct linear factors of $v(x)$ will be ti. Equating this number to the number of factors in $v^*(x)$, we obtain $ti = nr$, which implies $i = rs$ (using $n = st$). Thus,

$$P_i(h) \supseteq P(f) \cap P_r(g), \qquad i = rs, \quad r = 1, 2, \cdots, m. \tag{18}$$

We next prove the inclusion in (18) can be reversed and that $P_i(h) = \varnothing$ if $s \nmid i$. Suppose $p \in P_i(h)$ for some i, $1 \leqq i \leqq ms$. Since K and L are subfields of M, by Theorem 2, $P_i(h) \subseteq P(f) \cap P(g)$. Also, each $p \in P_i(h)$, with a finite number of exceptions, belongs to some $P_r(g)$, $1 \leqq r \leqq m$. If $p \in P_r(g)$, then since f splits completely (mod p), we again find that v^* has nr distinct linear factors, while v has ti linear factors. Thus, $r = i/s$. This shows r is uniquely determined by i, and when $s \nmid i$, at most a finite number of primes belong to $P_i(h)$. Also, when i is a multiple of s, then that multiple must be r, which implies $P_i(h) \subseteq P(f) \cap P_r(g)$.

Finally,

$$P(h) = \bigcup_{i=1}^{ms} P_i(h) = \bigcup_{r=1}^{m} [P(f) \cap P_r(g)] = P(f) \cap \left[\bigcup_{r=1}^{m} P_r(g) \right] = P(f) \cap P(g). \blacksquare$$

7. Concluding remarks. In the preceding sections we have obtained information about the prime divisors of polynomials from properties of their roots and associated fields. We now consider the converse problem: Do the prime divisors of polynomials determine their algebraic properties? This question was first investigated in 1880 by Kronecker [**11**], who laid the foundations for much of the research that has followed.

In our brief discussion of this question, we shall limit ourselves to considering true converses of Theorems 2 and 3, which represent quite well the results in this field. We shall present these converses without proof, however, since all the known proofs employ ideal theoretic and transcendental methods, which lie outside the scope of this paper. The reader who would like further information on these matters is referred to the rather full discussion in Hasse [**9**, vol. 2, pp. 138–146], where the basic references can be found.

We begin with the converse of Theorem 3: If two irreducible polynomials f and g of degree n are such that $P_i(f) = P_i(g)$, $i = 1, 2, \cdots, n$, then they belong

to the same field. That this statement is *false* was shown by Gassmann [7] in 1926, when he proved the existence of two 180th degree polynomials (!) belonging to *nonconjugate* fields, which not only satisfy the required conditions, but also factor in exactly the same way (mod p) for almost all primes p. (See also Schinzel [**19**].) One of the authors [**8**, p. 138] has recently given a pair of polynomials $f(x)=x^8-3\cdot 2^4$, $g(x)=x^8-3^7$, which has the Gassmann property and which provides a simpler counterexample to the converse of Theorem 3.

We can obtain a *true* converse to Theorem 3, however, if we impose the further condition that f and g be normal (see Bauer [**1**]). (We note in this case that $P_n(f)=P(f)$ and $P_n(g)=P(g)$.)

If we modify the hypothesis of the converse of Theorem 2 in a similar way, we obtain the THEOREM: *If $f:K$ with f normal, $g:L$, and $P(f)\supseteq P(g)$, then $K\subseteq L$.* (See Bauer [**2**].)

As a particular application of this theorem in the case K is the nth cyclotomic field, we have the result: *A field L will contain K if and only if for $g:L$, $P(g)$, with a finite number of exceptions, contains only primes $p\equiv 1$* (mod n). (See Theorem 9.)

In closing, we would like to remark it would be of some interest if elementary proofs of the theorems of this section could be constructed.

References

1. M. Bauer, Über einen Satz von Kronecker, Arch. Math. und Physik, Bd. 6, 218–219.

2. ———, Zur Theorie der algebraischen Zahlkörper, Math. Ann., 77 (1916) 353–356.

3. G. Birkhoff and S. MacLane, A Survey of Modern Algebra, 3rd ed., Macmillan, New York, 1965, Ch. XIV.

4. Z. I. Borevich and I. R. Shafarevich, Number Theory, Academic Press, New York, 1966.

5. R. Dedekind, Über den Zusammenhang zwischen der Theorie der Ideale und der höheren Kongruenzen, Abh. Akad. Wiss. Göttingen, 23 (1878) 3–37. (Repr. in Vol. 1, pp. 202–232 of the Gesammelte Mathematische Werke, Chelsea, New York, 1969. The theorem in question is Thm. I, pp. 212–213.)

6. L. Fjellstedt, Bemerkungen über gleichzeitige Lösbarkeit von Kongruenzen, Ark. Mat., 3 (1955) 193–198.

7. F. Gassmann, Bemerkungen zur vorstehenden Arbeit von Hurwitz, Math. Z., 25 (1926) 665–675.

8. I. Gerst, On the theory of nth power residues and a conjecture of Kronecker, Acta. Arith., 17 (1970) 121–139.

9. H. Hasse, Bericht über neuere Untersuchungen und Probleme aus der Theorie der algebraischen Zahlkörper, 2nd ed., Physica, Würzburg, 1965.

10. ———, Über das Problem der Primzerlegung in Galoisschen Zahlkörpern, S.-B. Berlin. Math. Ges., (1951/52) 8–27.

11. L. Kronecker, Über die Irreducibilität von Gleichungen, Monatsb. Deutsch. Acad. Wiss. Berlin, (1880) 155–163.

12. S. Lang, Algebraic Numbers, Addison-Wesley, Reading, Mass., 1964.

13. H. B. Mann, Introduction to Algebraic Number Theory, The Ohio State University Press, Columbus, 1955.

14. T. Nagell, Introduction to Number Theory, 2nd ed., Chelsea, New York, 1964, pp. 164–167.

15. ———, Zahlentheoretische Notizen I, Ein Beitrag zur Theorie der höheren Kongruenzen, Videnskapsselskapets Skrifter I, Matem.-Naturv. Klasse No. 13, Kristiania, 1923, 3–6.

16. T. Nagell, Sur les diviseurs premiers des polynômes, Acta Arith., 15 (1969) 235–244.
17. H. Pollard, The Theory of Algebraic Numbers, MAA Carus Mathematical Monographs No. 9, Wiley, New York, 1950.
18. P. Samuel, Théorie Algébrique des Nombres, Hermann, Paris, 1967.
19. A. Schinzel, On a theorem of Bauer and some of its applications, Acta Arith., 11 (1966) 333–344.
20. I. Schur, Über die Existenz unendlich vieler Primzahlen in einiger speziellen arithmetischen Progressionen, S.-B. Berlin. Math. Ges., 11 (1912) 40–50.
21. B. van der Waerden, Modern Algebra, Ungar, New York, 1953, vol. 1, Ch. 4 and 5.

SOME PROPERTIES OF COMPLETELY MULTIPLICATIVE ARITHMETICAL FUNCTIONS

T. M. APOSTOL, California Institute of Technology

An arithmetical function f which is not identically zero is called *multiplicative* if

$$(1) \qquad f(mn) = f(m)f(n) \qquad \text{whenever } (m, n) = 1,$$

and it is called *completely multiplicative* if

$$(2) \qquad f(mn) = f(m)f(n) \qquad \text{for all } m \text{ and } n.$$

If f is multiplicative, then $f(1)=1$ and

$$f(p_1^{a_1} \cdots p_s^{a_s}) = f(p_1^{a_1}) \cdots f(p_s^{a_s})$$

if the p_i are distinct primes.

This note discusses some properties of multiplicative and completely multiplicative functions which do not seem to be widely known but which seem appropriate for presentation in a course in elementary number theory. In particular, we discuss various necessary and sufficient conditions for a multiplicative function f to be completely multiplicative. From the definitions in (1) and (2) we see immediately that one such condition is that

$$(3) \qquad f(mn) = f(m)f(n) \qquad \text{whenever } (m, n) > 1;$$

another is that

$$(4) \qquad f(p^a) = f(p)^a \qquad \text{for all primes } p \text{ and all integers } a \geqq 1;$$

and yet another is that

$$(5) \qquad f(p^a) = f(p)f(p^{a-1}) \qquad \text{for primes } p \text{ and all integers } a \geqq 1.$$

Further necessary and sufficient conditions can be given in terms of the Dirichlet product (or convolution) $f * g$ which is defined by the relation

$$(f * g)(n) = - \sum_{d|n} f(d)g(n/d).$$

It is well known (see [1], [2], or [3]) that this operation is commutative and associative. Moreover, the arithmetical function I given by the formula

$$I(n) = \left[\frac{1}{n}\right] = \begin{cases} 1 & \text{if } n = 1, \\ 0 & \text{if } n > 1, \end{cases}$$

is an identity element for Dirichlet multiplication. Also, if $f(1) \neq 0$, then f has a unique inverse f^{-1} such that

$$f * f^{-1} = I. \tag{6}$$

In fact, f^{-1} can be computed from the recursion formulas

$$f^{-1}(1) = \frac{1}{f(1)}, \qquad f^{-1}(n) = -\frac{1}{f(1)} \sum_{\substack{d \mid n \\ d>1}} f(d) f^{-1}\left(\frac{n}{d}\right), \qquad n > 1.$$

Thus, the set of arithmetical functions with $f(1) \neq 0$ is an abelian group under Dirichlet multiplication. The multiplicative functions are a subgroup of this group. However, the completely multiplicative functions are not a subgroup since the Dirichlet product of two completely multiplicative functions need not be completely multiplicative.

We denote by u the arithmetical function which is identically 1,

$$u(n) = 1 \qquad \text{for all } n,$$

and we call this the *unit function*. Its Dirichlet inverse is the Möbius function μ. Thus, $\mu * u = I$, or

$$\sum_{d \mid n} \mu(d) = I(n).$$

The Möbius inversion formula simply states that the two equations

$$G = g * u \quad \text{and} \quad g = G * \mu$$

are equivalent.

The Möbius function has the further properties that $\mu(n) = 0$ if n is divisible by the square of a prime, whereas $\mu(p_1 \cdots p_s) = (-1)^s$ if the p_i are distinct primes.

Our first theorem relates the Dirichlet inverse f^{-1} with f and μ when f is multiplicative.

THEOREM 1. *Assume f is multiplicative. Then for every squarefree integer n we have*

$$f^{-1}(n) = \mu(n) f(n). \tag{7}$$

Moreover, if p is any prime we have

$$f^{-1}(p^2) = f(p)^2 - f(p^2). \tag{8}$$

Proof. Equation (7) holds trivially if $n=1$. If $n>1$ and n is squarefree then $n=p_1 \cdots p_s$, where the p_i are distinct primes. For each prime p, equation (6) gives us $f(1)f^{-1}(p)+f(p)f^{-1}(1)=0$, or

$$f^{-1}(p) = -f(p). \tag{9}$$

Therefore, since both f and f^{-1} are multiplicative, we have

$$f^{-1}(p_1 \cdots p_s) = (-1)^s f(p_1 \cdots p_s).$$

Since $\mu(p_1 \cdots p_s)=(-1)^s$, this proves (7). To prove (8) we take $n=p^2$ in (6) and use (9).

Next we prove that relation (7) is necessary and sufficient for a multiplicative function to be completely multiplicative.

THEOREM 2. *Assume f is multiplicative. Then f is completely multiplicative if and only if $f^{-1}(n)=\mu(n)f(n)$ for all n.*

Proof. Let $g(n)=\mu(n)f(n)$. If f is completely multiplicative we have

$$\sum_{d\mid n} g(d)f(n/d) = \sum_{d\mid n} \mu(d)f(d)f(n/d) = f(n)\sum_{d\mid n}\mu(d) = f(n)I(n).$$

But $f(n)I(n)=I(n)$ since $f(1)=1$, so $g=f^{-1}$.

Conversely, assume $f^{-1}(n)=\mu(n)f(n)$. To show that f is completely multiplicative it suffices to prove that $f(p^a)=f(p)f(p^{a-1})$ for all prime powers. Since $f^{-1}(n)=\mu(n)f(n)$ we have

$$\sum_{d\mid n}\mu(d)f(d)f(n/d) = 0 \qquad \text{for all } n>1.$$

Taking $n=p^a$ in this formula we obtain $\mu(1)f(1)f(p^a)+\mu(p)f(p)f(p^{a-1})=0$, which becomes $f(p^a)=f(p)f(p^{a-1})$. This implies that f is completely multiplicative.

As a consequence of Theorems 1 and 2 we obtain the following necessary and sufficient condition for a multiplicative function to be completely multiplicative.

THEOREM 3. *Assume f is multiplicative. Then f is completely multiplicative if and only if*

$$f^{-1}(p^a) = 0 \tag{10}$$

for all primes p and all integers $a \geqq 2$.

Proof. If f is completely multiplicative, Theorem 2 shows that $f^{-1}(p^a)=0$ if $a\geqq 2$. To prove the converse, we assume (10) and prove that $f^{-1}(n)=0$ whenever n has a square prime factor. Each such n has the form $n=p^a q$ where $a\geqq 2$ and $(q,p)=1$. Since f^{-1} is multiplicative we have

$$f^{-1}(n) = f^{-1}(p^a)f^{-1}(q) = 0.$$

Hence, by Theorem 1, $f^{-1}(n)=\mu(n)f(n)$ for all n so, by Theorem 2, f is com-

pletely multiplicative.

The next theorem gives a characterization of multiplicative functions in terms of a distributive property of Dirichlet multiplication,

$$f\cdot(g * h) = (f\cdot g) * (f\cdot h),$$

where the dot denotes ordinary multiplication of functions.

THEOREM 4. (a) *If f is completely multiplicative then*

$$f\cdot(g * h) = (f\cdot g) * (f\cdot h) \tag{11}$$

for all arithmetical functions g and h.

(b) *If f is multiplicative and if* (11) *holds for $g=\mu$ and $h=\mu^{-1}$, then f is completely multiplicative.*

Proof of (a). Assume f is completely multiplicative. Then $f(d)f(n/d)=f(n)$ whenever $d\mid n$, so we have

$$f(n)\sum_{d\mid n} g(d)h(n/d) = \sum_{d\mid n} f(d)g(d)f(n/d)h(n/d),$$

which proves (11).

Proof of (b). First we note that $f(n)I(n)=I(n)$ since $f(1)=1$. Taking $g=\mu$ and $h=\mu^{-1}$ in (11) we obtain

$$I(n) = \sum_{d\mid n} f(d)\mu(d)f(n/d),$$

which states that $f^{-1}(n)=\mu(n)f(n)$. Therefore, by Theorem 2, f is completely multiplicative.

As a corollary of Theorem 4 we have:

THEOREM 5. *Assume f is multiplicative. Then f is completely multiplicative if and only if $f\cdot(g * h)=(f\cdot g) * (f\cdot h)$ for all arithmetical functions g and h.*

We can also use Theorem 4 to prove:

THEOREM 6. (a) *If f is completely multiplicative, then we have*

$$(f\cdot g)^{-1} = f\cdot g^{-1} \tag{12}$$

for every arithmetical function g with $g(1)\neq 0$.

(b) *If f is multiplicative and if* (12) *holds for $g=u=\mu^{-1}$, then f is completely multiplicative.*

Proof of (a). Since f is completely multiplicative we can take $h=g^{-1}$ in Theorem 4 to obtain

$$f\cdot I = (f\cdot g) * (f\cdot g^{-1}).$$

Since $f\cdot I=I$ this proves that $(f\cdot g)^{-1}=f\cdot g^{-1}$.

Proof of (b). Taking g to be the unit function in (12) we have $f\cdot g=f$, and

(12) becomes $f^{-1}=f\cdot\mu$. By Theorem 2 this proves that f is completely multiplicative.

As a corollary of Theorem 6 we have:

THEOREM 7. *Assume f is multiplicative. Then f is completely multiplicative if and only if*

$$(f\cdot g)^{-1} = f\cdot g^{-1}$$

for every arithmetical function g with $g(1)\neq 0$.

In a recent issue of this MONTHLY, R. Sivaramakrishnan [4] has offered the following theorem as Problem E 2196:

If f is multiplicative, then f is completely multiplicative if and only if

$$\sum_{d\mid n} f(d)f^{-1}(n/d)d = f(n)\phi(n),$$

where $\phi(n)$ is the Euler totient function.

Since $\phi(n)=\sum_{d\mid n} d\mu(n/d)$, this result is a special case of the following theorem:

THEOREM 8. *Let G be any arithmetical function and let*

$$g(n) = \sum_{d\mid n} G(d)\mu(n/d). \tag{13}$$

(a) *If f is completely multiplicative, then for every G we have*

$$\sum_{d\mid n} f(d)f^{-1}(n/d)G(d) = f(n)g(n). \tag{14}$$

(b) *Assume f is multiplicative. If* (14) *holds for some completely multiplicative G such that $G(p)\neq 1$ for all primes p, then f is completely multiplicative.*

Proof of (a). Equation (14) states that $(f\cdot G) * f^{-1}=f\cdot g$. If f is completely multiplicative we have $f^{-1}=f\cdot\mu$. Since $g=G*\mu$ we are to prove that $(f\cdot G)*(f\cdot\mu)=f\cdot(G*\mu)$. But this follows at once from Theorem 4(a).

Proof of (b). We are given that f is multiplicative and that (14) holds for some completely multiplicative G with $G(p)\neq 1$ for all primes p. We shall show that f is completely multiplicative by proving that $f^{-1}(p^a)=0$ for all primes p and all integers $a\geqq 2$.

The proof is by induction on a. For the case $a=2$ we take $n=p^2$ in (14) and obtain

$$f^{-1}(p^2) + G(p)f(p)f^{-1}(p) + G(p^2)f(p^2) = f(p^2)g(p^2). \tag{15}$$

From Equation (7) in Theorem 1 we have $f^{-1}(p)=-f(p)$. Also, (13) gives us $g(p^2)=G(p^2)-G(p)$, so (15) can be written as

$$f^{-1}(p^2) = G(p)[f(p)^2 - f(p^2)]. \tag{16}$$

But from Equation (8) in Theorem 1 we also have $f(p)^2-f(p^2)=f^{-1}(p^2)$, so (16) becomes

$$f^{-1}(p^2) = G(p)f^{-1}(p^2).$$

Since $G(p)\neq 1$ this implies $f^{-1}(p^2)=0$.

Now we make the induction hypothesis that $f^{-1}(p^t)=0$ for every integer t in the range $2\leqq t\leqq a-1$, where $a\geqq 3$. Taking $n=p^a$ in (14) we find

$$f^{-1}(p^a) + G(p^{a-1})f(p^{a-1})f^{-1}(p) + G(p^a)f(p^a) = f(p^a)g(p^a), \tag{17}$$

the remaining terms in the sum being zero. Using the relation $f^{-1}(p)=-f(p)$ again and the relation $g(p^a)=G(p^a)-G(p^{a-1})$ we can rewrite (17) as

$$f^{-1}(p^a) = G(p^{a-1})[f(p)f(p^{a-1}) - f(p^a)]. \tag{18}$$

But if we take $n=p^a$ in the identity

$$\sum_{d|n} f(d)f^{-1}(n/d) = 0$$

and again use the induction hypothesis we find $f(p)f(p^{a-1})-f(p^a)=f^{-1}(p^a)$. Using this in (18) we get $f^{-1}(p^a)=G(p)^{a-1}f^{-1}(p^a)$ which implies $f^{-1}(p^a)=0$ since $G(p)\neq 1$. This proves, by induction, that $f^{-1}(p^a)=0$ for every prime p and all integers $a\geqq 2$. Therefore, by Theorem 3, f is completely multiplicative.

References

1. T. M. Apostol, A characteristic property of the Möbius function, this MONTHLY, 72 (1965) 279–282.
2. E. T. Bell, Outline of a theory of arithmetical functions in their algebraic aspects, J. Indian Math. Soc., 17 (1927–28) 249–260.
3. J. Popken, On convolutions in prime number theory, Nederl. Akad. Wetensch. Proc. Ser. A 58(1955) 10–15.
4. R. Sivaramakrishnan, Problem E 2196, this MONTHLY, 76 (1969) 1062.

MATHEMATICAL NOTES

EDITED BY DAVID DRASIN

Manuscripts for this Department should be sent to David Drasin, Division of Mathematical Sciences, Purdue University, Lafayette, IN 47907.

ON THE SUM OF THE kth POWERS OF THE FIRST n INTEGERS

J. L. PAUL, Purdue University (Now at the University of Cincinnati)

Set $S_k(n)=1^k+2^k+\cdots+n^k$, where k is a non-negative integer. The identity

$$1 + \sum_{k=0}^{r-1}\binom{r}{k}S_k(n) = (n+1)^r \tag{1}$$

is given on page 160 of J. Riordan, *Combinatorial Identities* (Wiley, 1968) with a proof involving generating functions. I shall give an elementary combinatorial proof.

Consider the "cube" C of integer lattice points $(x_1, x_2, \cdots, x_r)$, $0 \leqq x_j \leqq n$. There are $(n+1)^r$ points in C.

Now partition C: for each k, let C_k be the set of points of C for which exactly $r-k$ coordinates have a common value and the remaining k coordinates are strictly smaller. Each point of C falls into exactly one set C_k, hence

$$(n+1)^r = |C| = \sum_{k=0}^{r-1} |C_k|, \tag{2}$$

where $|C|$ denotes the number of points in C. Thus to prove (1) it suffices to combine (2) with the following two formulas:

$$|C_k| = \binom{r}{k} S_k(n) \qquad (1 \leqq k \leqq r-1), \tag{3}$$

$$|C_0| = 1 + n = 1 + \binom{r}{0} S_0(n). \tag{4}$$

Relation (4) is obvious, because C_0 consists of the $n+1$ diagonal points $(x, x, \cdots, x)$ for $0 \leqq x \leqq n$. To prove (3), first note that there are

$$\binom{r}{r-k} = \binom{r}{k}$$

ways to pick the $r-k$ coordinates having the largest value; this largest value L ranges from 1 to n. Having picked these largest coordinates, say for definiteness $x_{k+1} = x_{k+2} = \cdots = x_r = L$, then $0 \leqq x_1 \leqq L-1, \cdots, 0 \leqq x_k \leqq L-1$, so there are L^k corresponding points $(x_1, \cdots, x_k, L, \cdots, L)$. Since $1 \leqq L \leqq n$, there are a total of $1^k + 2^k + \cdots + n^k = S_k(n)$ points in C_k with the particular $r-k$ largest coordinates chosen. Thus, all together, $|C_k| = \binom{r}{k} S_k(n)$, which completes the proof of (3), and we are done.

A PROOF OF THE DIVERGENCE OF $\Sigma\, 1/p$

IVAN NIVEN, University of Oregon

First we prove that $\Sigma'\, 1/k$ diverges, where Σ' denotes the sum over the squarefree positive integers. Each positive integer is uniquely expressible as a product of a squarefree positive integer and a square, so for any positive integer n,

$$\left(\sum_{k<n}{}' \, 1/k\right)\left(\sum_{j<n} 1/j^2\right) \geqq \sum_{m<n} 1/m.$$

Here the second sum is bounded but the third sum is unbounded as n increases, so the first sum must be unbounded. Next suppose that $\Sigma\ 1/p$ converges to β, the sum taken over all primes p. By dropping all terms beyond x in the series expansion of e^x or $\exp(x)$, we see that $\exp(x) > 1+x$ for $x>0$. Hence for each positive integer n

$$\exp(\beta) > \exp\left(\sum_{p<n} 1/p\right) = \prod_{p<n} \exp(1/p) > \prod_{p<n} (1 + 1/p) \geqq \sum_{k<n}{}' 1/k.$$

But this contradicts the unboundedness of the last sum, so $\Sigma\ 1/p$ diverges.

AN ALTERNATIVE PROOF OF A THEOREM OF ERDÖS AND SZEKERES

PAUL BLACKWELL, University of Missouri

Using maximal length decreasing sequences, Seidenberg [2] gave an elegant indirect proof of a theorem of Erdös and Szekeres which was proved in [1] by an ingenious inductive argument. The short direct proof given here employs maximal leftmost decreasing sequences. Let $S=\{s_1, s_2, \cdots, s_r\}$ be a sequence of distinct real numbers. A monotone decreasing subsequence S' will be called **leftmost** if $s_1'=s_1$ and each term s_i' of S' is the next term in S which is smaller than s_{i-1}'. A leftmost decreasing subsequence is maximal if no leftmost decreasing subsequence is longer.

THEOREM. *Let $S=\{s_1, s_2, \cdots, s_r\}$ be a sequence of $r>mn$ distinct real numbers. If each decreasing subsequence of S has at most m terms, there exists an increasing subsequence of S with more than n terms.*

Proof. Successively remove from S the maximal leftmost decreasing subsequences $S_1, S_2, \cdots, S_t$ until S is exhausted. Then $t>n$, because each term in S appears in exactly one subsequence S_i and, by hypothesis, each such subsequence has at most m terms. For each term s of S_k $(1<k\leqq t)$ there is a term of S_{k-1} that precedes it in S and is smaller, for otherwise s would have been chosen as a term of S_{k-1}. The desired monotone increasing subsequence $\{a_1, a_2, \cdots, a_t\}$ of S can therefore be chosen as follows: Choose a_t arbitrarily from S_t, choose a_{t-1} as any term in S_{t-1} that both precedes a_t and is smaller, and so on.

References

1. P. Erdös and G. Szekeres, A combinatorial problem in geometry, Compositio Math., 2 (1935) 463–470.

2. A. Seidenberg, A simple proof of a theorem of Erdös and Szekeres, J. London Math., Soc., 34 (1959) 352.

RESEARCH PROBLEMS

EDITED BY RICHARD GUY

In this Department the Monthly presents easily stated research problems dealing with notions ordinarily encountered in undergraduate mathematics. Each problem should be accompanied by relevant references (if any are known to the author) and by a brief description of known partial results. Manuscripts should be sent to Richard Guy, Department of Mathematics, Statistics, and Computing Science, The University of Calgary, Calgary 44, Alberta, Canada.

AN EQUATION FOR FINITE GROUPS

ROBERT HIGGINS AND DAVID BALLEW, South Dakota School of Mines and Technology

Let G be a finite group of order g. The (conjugate) class of x in G is the set of all multiples $y^{-1}xy$, where y runs through the elements of G. A class is **ambivalent** if x^{-1} is in the class whenever x is. A group is ambivalent if all of its classes are ambivalent. Let $\rho_n(a)$ represent the number of solutions to the equation $x^n = a$ in G for a in G.

The following theorem is known [2, p. 146], but we include an elementary proof for completeness and motivation for our subsequent comments.

THEOREM. *Let G be a finite group of order g with α ambivalent classes. Then $\sum_{a\in G}(\rho_2(a))^2 = \alpha g$, where the sum runs over all elements a of G.*

Proof. Since $\rho_2(a) = 0$ for all elements of G which are not squares,

$$\sum_{a\in G}(\rho_2(a))^2 = \sum_{b^2\in G}(\rho_2(b^2))^2.$$

Now focus on some particular b^2 in G. Suppose $b_1, b_2, b_3, \cdots, b_n$ are the n elements in G such that $b_i^2 = b^2$. Clearly, we have

$$(\rho_2(b^2))^2 = \sum_{b_i^2=b_j^2} 1 \qquad \text{where } 1 \leqq i \leqq n, \quad 1 \leqq j \leqq n,$$

since $(\rho_2(b^2))^2 = n^2$ and there are n^2 ways to have $b_i^2 = b_j^2$. Therefore, if we consider $\sum_{x^2=y^2} 1$ for all unordered pairs (x, y) such that $x^2 = y^2$, we must obtain $\sum_{b^2\in G}(\rho_2(b^2))^2$. Hence the problem remaining is to determine all pairs (x, y) such that $x^2 = y^2$.

Let $y = cx$; we shall count all pairs (x, cx) such that $x^2 = (cx)^2$ or $x^2 = cxcx$; that is, $c = xc^{-1}x^{-1}$. This means that c belongs to some ambivalence class. For each $c \in G$, the number of x such that $c = xc^{-1}x^{-1}$ is the index I_c in G of the class to which c belongs. Therefore the number of possible pairs (x, y) is $\sum_{c \text{ ambivalent}} I_c$, and by our previous statements

$$\sum_{a\in G}(\rho_2(a))^2 = \sum_{c \text{ ambivalent}} I_c.$$

But, by Lagrange's theorem, I_c summed over each element in one class is g. Hence summing over all ambivalent classes, $\sum_{c \text{ ambivalent}} I_c = \alpha g$, where α is the number of ambivalent classes.

We shall now state the basic conjecture of this paper:

CONJECTURE. Let G be a finite group of order g. Then for integers n and k there is an integer $\beta = \beta(n, k)$ such that

$$\sum_{a \in G} (\rho_n(a))^k = \beta g.$$

This conjecture has been tested by the authors on a large number of finite groups for several values of n and k and no counterexample has been found, [3].

There are several special cases of the conjecture that can be handled quite easily. For example, one can show that if the exponent of a group is j and $(n, j) = m$, then $\rho_n(a) = \rho_m(a)$ for all a in G. Further, one can show that if the conjecture holds for a collection of groups, then it holds for their direct product. Thus since the conjecture is true for cyclic groups, it is true for all finite Abelian groups.

There are very few theorems involving solutions of equations in finite groups. Perhaps the most famous are those due to Frobenius and Philip Hall [1, pp. 136–137]. However, the known results do not seem to apply to our conjecture. Furthermore, neither our proof of the case $n = k = 2$ nor the other known proofs seem to generalize. The roadblock appears to be a generalization of the ambivalence class idea, and it will be interesting to see, when the theorem is proved, what the β in the conjecture turns out to be.

References

1. Marshall Hall, The Theory of Groups, Macmillan, New York, 1959.
2. Morton Hamermesh, Group Theory, Addison-Wesley, New York, 1962.
3. Robert Higgins, Numerical relationships in group theory, M.S. Thesis, South Dakota School of Mines and Technology, Rapid City, 1970.

A PROBLEM ON INEQUALITIES

FRANZ HERING, University of Washington, Seattle

A finite set I of natural numbers is said to be *alternating*, provided that there is an odd member of I between any two even members, and an even member between any two odd members; equivalently, the odd and even members alternate when I is arranged as an increasing sequence. For $0 \leqq r \leqq m$ let $\mathfrak{A}_{r,m}$ be the set of all alternating subsets of $\{1, 2, \cdots, m\}$ with r elements. For example,

$$\mathfrak{A}_{3,5} = \{\{1, 2, 3\}, \{1, 2, 5\}, \{1, 4, 5\}, \{2, 3, 4\}, \{3, 4, 5\}\}.$$

We are concerned with the behaviour of the *alternating polynomial*

$$A_r(x_1, \cdots, x_m) = \sum \left\{ \prod_{i \in I} x_i : I \in \mathfrak{A}_{r,m} \right\}$$

on the simplex

$$S_m = \{(x_1, \cdots, x_m) \in R^m : x_i \geqq 0 \text{ for } i = 1, \cdots, m, \sum x_i = 1\}.$$

Let $\beta_{r,m}$ denote the maximum of A_r on S_m. Then we conjecture that

$$\beta_{r,m} = \begin{cases} A_r\left(\frac{1}{m}, \cdots, \frac{1}{m}\right) & \text{for } r \equiv m \pmod 2 \\ A_r\left(\frac{1}{m-1}, \cdots, \frac{1}{m-1}, 0\right) & \text{for } r \not\equiv m \pmod 2. \end{cases}$$

This implies, in particular, that the maximum $\beta_{r,m}$ is attained at an interior point of S_m when $r \equiv m \pmod 2$, and at a boundary point when $r \not\equiv m \pmod 2$. Furthermore, we conjecture that $\beta_{r,m}$ is attained *uniquely* at $(1/m, \cdots, 1/m)$ if and only if $r \equiv m \pmod 2$ and $r > 2$. We proved these conjectures for $r \leqq 4$ in [1]. When $r = m$, the fact that $\beta_{m,m} = A_m(1/m, \cdots, 1/m) = m^{-m}$ is equivalent to the arithmetic-geometric mean inequality.

Reference

1. Franz Hering, A generalisation of the arithmetic-geometric mean inequality and an application to finite sequences of zeros and ones (to appear in the Israel J. Math.).

CLASSROOM NOTES

EDITED BY DAVID DRASIN

Manuscripts for this Department should be sent to David Drasin, Division of Mathematical Sciences, Purdue University, Lafayette, IN 47907.

A COMBINATORIAL PROOF OF A PARTITION FUNCTION LIMIT

G. E. ANDREWS, Pennsylvania State University

If $p(n)$ denotes the ordinary partition function, the result

$$\lim_{n\to\infty} p(n)^{1/n} = 1$$

has both theoretical and pedagogical interest.

Most proofs of this result rely in one way or another on function theoretic properties of the generating function of $p(n)$ [3; p. 208], [2; pp. 274–276]. Although elementary arguments have been used to estimate the size of $p(n)$ (e.g., [1]), none seems really appropriate in an elementary introduction to partitions.

If a simple direct proof of $\lim_{n\to\infty} p(n)^{1/n} = 1$ can be given, then the root test gives the radius of convergence of the power series $\sum_{n=0}^{\infty} p(n)x^n$ as 1, and this knowledge simplifies further rigorous developments of the generating function. We prove the theorem by direct combinatorial considerations.

DEFINITION. The symbol $p_k(n)$ denotes the number of partitions of n into at most k parts.

LEMMA 1. *The relation* $p_k(n) \leqq (n+1)^k$ *holds for each integer* $k > 0$.

Proof. If in the expression $a_1 + a_2 + \cdots + a_k$ we allow each a to take all integral values in the interval $[0, n]$, we obtain all partitions of n into k parts as well as many partitions of other numbers. Since there are $(n+1)^k$ ways of making such substitutions the lemma follows.

LEMMA 2. *The relation* $p(n) \leqq p(n-1) + p_k(n) + p(n-k)$ *holds for each integer* $k > 0$.

Proof. Separate the partitions of n into 3 classes: (1) those partitions which contain 1 as a summand, (2) those partitions which contain no 1's and have at most k parts, and (3) those partitions which contain no 1's and have more than k parts.

By deleting a 1 from each element of the first class we see that there are exactly $p(n-1)$ elements in this class. The second class clearly contains at most $p_k(n)$ elements. In the third class subtract 1 from the smallest k summands of each partition; this establishes a one-to-one correspondence between the elements of the third class and a subset of the partitions of $n-k$. Hence the third class has at most $p(n-k)$ elements. Consequently $p(n) \leqq p(n-1) + p_k(n) + p(n-k)$, as desired.

THEOREM. $\text{Lim}_{n\to\infty}\, p(n)^{1/n} = 1$.

Proof. It is sufficient to establish that for any $\epsilon > 0$,

$$p(n) < K(1+\epsilon)^n \quad \text{for } n \text{ sufficiently large.} \tag{1}$$

Given (1), we deduce directly that $1 \leqq p(n)^{1/n} < K^{1/n}(1+\epsilon) \to 1+\epsilon$ as $n \to \infty$.

First we choose k sufficiently large so that $(1+\epsilon)^{k-1} > 2/\epsilon$. Next by Lemma 1, we choose n_0 so large that for $n > n_0$

$$p_k(n) < \frac{\epsilon}{2}(1+\epsilon)^{n-1};$$

this is possible since $p_k(n) \leqq (n+1)^k$ and $\lim_{n\to\infty} [(n+1)^k/(1+\epsilon)^{n-1}] = 0$ by a k-fold application of L'Hospital's Rule.

Now let

$$K = \left(\max_{0 \leqq n \leqq n_0} \frac{p(n)}{(1+\epsilon)^n}\right) + 1.$$

Then $p(n) < K(1+\epsilon)^n$, for all $n \leqq n_0$.

Assume $p(n) < K(1+\epsilon)^n$ for all $n < m$, where $m > n_0$. Then

$$p(m) \leqq p(m-1) + p_k(m) + p(m-k)$$

$$< K(1+\epsilon)^{m-1} + \frac{\epsilon}{2}(1+\epsilon)^{m-1} + K(1+\epsilon)^{m-k}$$

$$< K(1+\epsilon)^{m-1}\left(1+\frac{\epsilon}{2}+\frac{1}{(1+\epsilon)^{k-1}}\right)$$

$$< K(1+\epsilon)^{m-1}\left(1+\frac{\epsilon}{2}+\frac{\epsilon}{2}\right) = K(1+\epsilon)^{m}.$$

Hence, by mathematical induction, (1) is valid for all sufficiently large n.

Partially supported by National Science Foundation Grant GP-9660.

References

1. P. Erdös, On an elementary proof of some asymptotic formulas in the theory of partitions, Ann. of Math., 43 (1942) 437–450.

2. G. H. Hardy and E. M. Wright, An Introduction to the Theory of Numbers, Oxford University Press, Oxford, 4th ed., 1960.

3. I. Niven and H. S. Zuckerman, An Introduction to the Theory of Numbers, Wiley, New York, 1960.

ON THE CONVERSE HELLY THEOREM

H. GUGGENHEIMER, Polytechnic Institute of Brooklyn

Helly's theorem is one of the fundamental results in convexity (for a comprehensive survey, see [1]). The importance of Helly's theorem would make the discussion of a converse a desirable topic for an undergraduate introduction to convexity. There exists a converse Helly theorem due to Dvoretzky [2], but its proof depends on the Krein-Milman theorem and is unsuitable for undergraduate presentation. When I was last teaching undergraduate convexity, I found that Dvoretzky's argument hides a simpler theorem that needs only tools similar to those used for the proof of Helly's theorem itself.

The Helly property of a set $\mathcal{C}$ of sets in n-space E^n is:

For $m \geqq n+1$, either the intersection of m of the sets of $\mathcal{C}$ is not empty, or the intersection of some $n+1$ among the m sets is empty.

Helly's theorem states that the finite sets of convex sets and the sets of convex, compact sets are sets $\mathcal{C}$ with the Helly property.

It is not true that the members of a set $\mathcal{C}$ with the Helly property in E^n are necessarily convex; it is easy to build finite sets $\mathcal{C}$ of nonconvex sets with the Helly property and there exists a topological Helly theorem. Dvoretzky's theorem shows that these cases are accidents of special position: If we add to any given $\mathcal{C}$ all sets obtained from the elements of $\mathcal{C}$ by moving them around and squishing them, then the Helly property is conserved if and only if the sets of $\mathcal{C}$ are convex. In a more mathematical language: *Let $\mathcal{AC}$ be the set of all images of*

the elements of $\mathcal{C}$ *by affine transformations of nonzero determinant. If all sets of* $\mathcal{C}$ *are n-dimensional, then the sets of* $\mathcal{C}$ *are convex if and only if* $\mathcal{AC}$ *has the Helly property.* We use the same basic idea: If $\mathcal{C}$ is embedded in a sufficiently huge set and the Helly property is conserved, then the elements of $\mathcal{C}$ must be convex.

We denote by $\mathcal{S}$ the set of all nondegenerate simplices of E^n (the convex hulls of $n+1$-ples of points that are not in the same hyperplane) and by conv C the convex hull of a set C.

THEOREM. *Let* $\mathcal{C}$ *be a finite set of closed sets or a set of compact sets of* E^n *none of which is contained in a hyperplane. All sets of* $\mathcal{C}$ *are convex if and only if* $\mathcal{S}\cup\mathcal{C}$ *has the Helly property.*

We show that C is convex if $\mathcal{S}\cup\{C\}$ has the Helly property. The converse is Helly's theorem.

Since C is not in a hyperplane, conv C contains a nondegenerate simplex, and its interior, int conv $C\neq\emptyset$. By a theorem of Steinitz ([**3**], Theorem 3.13) each point $x\in$int conv C is in the interior of the convex hull of at most $2n$ points of C. The polytope spanned by these points can be divided into simplices by a finite number of $(n-1)$-dimensional faces. The union of these faces is a set of n-dimensional volume zero. Therefore, in the neighborhood of any point of int conv C there are points in the interior of some nondegenerate simplex spanned by the points of C. This proves the following complement to Carathéodory's theorem ([**3**], Theorem 1.20):

LEMMA. *The points in the interiors of the simplices spanned by the points of C form a dense set in* int conv C.

We show now that the Helly property for $\mathcal{S}\cup\{C\}$ implies that the points in the interiors of the simplices spanned by points of C are in C. For

$$x \in \text{int conv}\{x_0, \cdots, x_n\} \qquad x_i \in C, \quad i = 0, \cdots, n$$

we look at the simplices S_j (where $j=0, \cdots, n$) defined by one face and the vertex x:

$$S_j = \text{conv}\{x, x_0, \cdots, x_{j-1}, x_{j+1}, \cdots, x_n\}.$$

Then $S_0\cap S_1\cap \cdots \cap S_n=\{x\}$. The Helly property now implies $x\in C$.

By the lemma, $C\cap$int conv C is dense in int conv C. Since C is closed, C = conv C is a convex set.

References

1. L. Danzer, B. Grünbaum, V. Klee, Helly's theorem and its relatives. Convexity, Proceedings of Symposia in Pure Mathematics, VII, Amer. Math. Soc., 1963, pp. 101–180.

2. A. Dvoretzky, A converse of Helly's theorem on convex sets, Pac. J. Math., 5 (1955) 345–350.

3. F. A. Valentine, Convex Sets, McGraw-Hill, New York, 1964.

ELEMENTARY APPROXIMATIONS TO THE AREA OF N-DIMENSIONAL ELLIPSOIDS

M. S. KLAMKIN, Ford Scientific Laboratory

The purpose of this note is to present an elementary self-contained derivation of sharp bounds for the perimeter of an ellipse and the surface area of an ellipsoid. The method can be easily extended for ellipsoids of higher dimension.

The perimeter P of the ellipse $x=a\cos\theta$, $y=b\sin\theta$ $(a\geqq b)$ is given by

$$P=\oint\sqrt{dx^2+dy^2}=4\int_0^{\pi/2}\sqrt{a^2\sin^2\theta+b^2\cos^2\theta}\,d\theta. \tag{1}$$

Consequently, good bounds on P can be obtained if we can obtain good bounds on the integrand. To this end, we symmetrize P by writing it in the equivalent form

$$P=2\int_0^{\pi/2}F(a,b,\theta)d\theta,$$

where $F(a,b,\theta)=F_1+F_2$ and

$$F_1=\sqrt{a^2\sin^2\theta+b^2\cos^2\theta},\qquad F_2=\sqrt{a^2\cos^2\theta+b^2\sin^2\theta}. \tag{2}$$

The upper and lower bounds for F are given in Hobson [1] and are obtained by the following simple algebraic argument: If we let

$$\lambda=a^2\cos^2\theta+b^2\sin^2\theta=\frac{a^2+b^2}{2}+\frac{a^2-b^2}{2}\cos^2\theta,$$

then $F=\sqrt{\lambda}+\sqrt{a^2+b^2-\lambda}$ and

$$F^2=a^2+b^2+\sqrt{(a^2+b^2)^2-(a^2+b^2-2\lambda)^2}.$$

Thus F is a maximum when $2\lambda=a^2+b^2$ and is a minimum when $(a^2+b^2-2\lambda)^2$ is a maximum. The latter occurs for $\cos 2\theta=\pm 1$, whence

$$a+b\leqq F\leqq\sqrt{2(a^2+b^2)},\quad\text{and}\quad \pi(a+b)\leqq P\leqq\pi\sqrt{2(a^2+b^2)}.$$

The inequalities are sharp for $a=b$, since then $P=2\pi a$.

Before proceeding to an alternate derivation for the bounds of F which will lead to extensions in higher dimensions, it is of interest to give two other derivations for the same bounds on P even though they are not quite as elementary.

An upper bound can be obtained from (1) by an immediate application of the Schwarz-Buniakowsky inequality, i.e.,

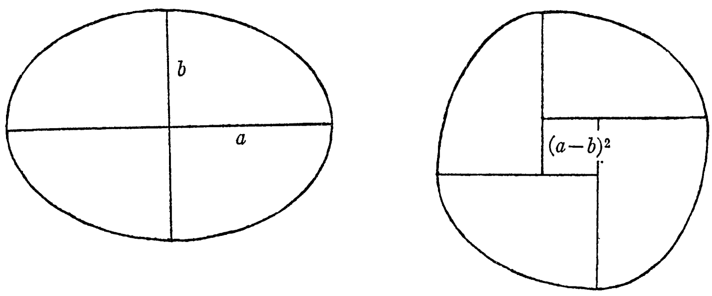
$$\int_n^m F^2(x)dx\cdot\int_n^m G^2(x)dx\geqq\left\{\int_n^m F(x)G(x)dx\right\}^2.$$

This gives

$$\int_0^{\pi/2} d\theta \int_0^{\pi/2} (a^2 \sin^2\theta + b^2 \cos^2\theta) d\theta \geqq \left\{ \int_0^{\pi/2} \sqrt{a^2 \sin^2\theta + b^2 \cos^2\theta}\, d\theta \right\}^2.$$

Integrating the l.h.s. yields the desired bound.

The lower bound has been obtained geometrically by Chakerian [2] using the isoperimetric inequality on the following rearrangement of an ellipse given in Steinhaus [3]:



By the isoperimetric inequality, the ratio of the square of the perimeter to the area of any figure is $\geqq$ the corresponding ratio for a circle, whence

$$\frac{P^2}{\pi ab + (a-b)^2} \geqq \frac{(2\pi r)^2}{\pi r^2} = 4\pi$$

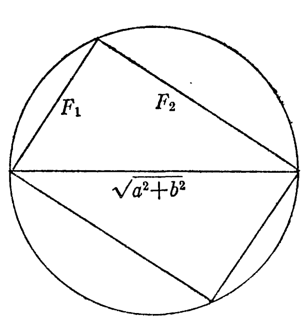
or

$$P^2 \geqq 4\pi\{(a-b)^2 + \pi ab\} \geqq \pi^2(a+b)^2 + \pi(4-\pi)(a-b)^2 \geqq \pi^2(a+b)^2.$$

If we applied the isoperimetric inequality directly to the ellipse, we would obtain the weaker inequality $P \geqq 2\pi\sqrt{ab}$.

From (2) it follows that

$$F_1^2 + F_2^2 = a^2 + b^2.$$



A geometrical interpretation for the bounds gotten on F corresponds to determining the rectangle of maximum and minimum semi-perimeter inscribed in a

circle of diameter $\sqrt{a^2+b^2}$. One can show geometrically that the maximum occurs when the rectangle is a square. The minimum occurs for the least "square" rectangle subject to the condition $F_1, F_2 \geqq b$.

We now consider bounds for the area of an ellipsoid. Our method for obtaining them will be applicable with little change for n-dimensional ellipsoids.

The surface area S of the ellipsoid

$$\left(\frac{x}{a}\right)^2 + \left(\frac{y}{b}\right)^2 + \left(\frac{z}{c}\right)^2 = 1$$

is given by

$$S = 8\int_{x=0}^{a}\int_{y=0}^{b}\left\{1 + \left(\frac{\partial z}{\partial x}\right)^2 + \left(\frac{\partial z}{\partial y}\right)^2\right\}^{1/2} dxdy.$$

Now let $x = a\cos\alpha$, $y = b\cos\beta$, $z = c\cos\gamma$, where $\cos^2\alpha + \cos^2\beta + \cos^2\gamma = 1$ to give

$$(3)\quad S = 8\int_0^{\pi/2}\int_0^{\pi/2}\left\{b^2c^2\cos^2\alpha + c^2a^2\cos^2\beta + a^2b^2\cos^2\gamma\right\}^{1/2}\frac{\sin\alpha\sin\beta}{\cos\gamma}\,d\alpha d\beta.$$

By symmetry,

$$S = \frac{8}{3}\int_0^{\pi/2}\int_0^{\pi/2}(F_1 + F_2 + F_3)\frac{\sin\alpha\sin\beta}{\cos\gamma}\,d\alpha d\beta,$$

where

$$F_1^2 = r^2\cos^2\alpha + s^2\cos^2\beta + t^2\cos^2\gamma,$$

$$F_2^2 = r^2\cos^2\gamma + s^2\cos^2\alpha + t^2\cos^2\beta,$$

$$F_3^2 = r^2\cos^2\beta + s^2\cos^2\gamma + t^2\cos^2\alpha,$$

and $r = bc$, $s = ca$, $t = ab$.

Since for $a = b = c$, the ellipsoid is a sphere, (3) reduces to

$$\pi = 2a^2\int_0^{\pi/2}\int_0^{\pi/2}\frac{\sin\alpha\sin\beta}{\cos\gamma}\,d\alpha d\beta.$$

Consequently, it suffices to find good constant bounds on $(F_1+F_2+F_3)$ in order to obtain corresponding bounds on S. Since

$$F_1^2 + F_2^2 + F_3^2 = r^2 + s^2 + t^2,$$

an upper bound for $\sum F_i$ will correspond geometrically to finding the rectangular parallelepiped of maximum edge length inscribed in a sphere of diameter $\sqrt{r^2+s^2+t^2}$. Intuitively, we would expect the figure to be a cube. A proof of this follows immediately from the known elementary inequality

$$(F_1 + F_2 + \cdots + F_n)^2 \leqq n(F_1^2 + F_2^2 + \cdots + F_n^2), \tag{4}$$

where $F_i \geqq 0$ and equality only if all the F_i's are equal. A proof of the latter inequality is obtained by simply expanding out the l.h.s. and replacing the "$2F_iF_j$" terms by means of the inequality $2F_iF_j \leqq F_i^2 + F_j^2$. The upper bound is then

$$\sum F_i \leqq \sqrt{3(b^2c^2 + c^2a^2 + a^2b^2)}.$$

Another way to obtain this bound and inequality (4) is to note that the function $\sqrt{x}$ is concave. Thus

$$\frac{\sqrt{F_1^2} + \sqrt{F_2^2} + \sqrt{F_3^2}}{3} \leqq \sqrt{\frac{F_1^2 + F_2^2 + F_3^2}{3}}.$$

A nonsharp lower bound can be obtained by applying the triangle inequality to the skew polygon whose sides are F_1, F_2, F_3 and $\sqrt{r^2+s^2+t^2}$. This gives

$$F_1 + F_2 + F_3 \geqq \sqrt{r^2 + s^2 + t^2}.$$

However, with slightly more trouble, we can obtain a better lower bound. We apply the triangle inequality to the 3 vectors

$$\begin{aligned} \mathbf{F}_1 &= \langle r \cos \alpha, s \cos \beta, t \cos \gamma \rangle, \\ \mathbf{F}_2 &= \langle s \cos \alpha, t \cos \beta, r \cos \gamma \rangle, \\ \mathbf{F}_3 &= \langle t \cos \alpha, r \cos \beta, s \cos \gamma \rangle. \end{aligned}$$

Since $|\mathbf{F}_1| + |\mathbf{F}_2| + |\mathbf{F}_3| \geqq |\mathbf{F}_1 + \mathbf{F}_2 + \mathbf{F}_3|$, we get

$$F_1 + F_2 + F_3 \geqq r + s + t = bc + ca + ab.$$

The corresponding bounds for S are now

$$4\pi(bc + ca + ab)/3 \leqq S \leqq 4\pi\{(b^2c^2 + c^2a^2 + a^2b^2)/3\}^{1/2}.$$

It is to be noted that although the derivation here appears to be new, the bounds obtained are well known. For a class of better approximations and a discussion of the accuracy of all these approximations, see the papers of Polya [4], Lehmer [5] and Carlson [6].

References

1. E. W. Hobson, A Treatise on Plane and Advanced Trigonometry, Dover, New York, 1957, p. 87.
2. G. D. Chakerian, On estimating the perimeter of an ellipse, Elem. der Math., Vol. XX/4 (1965).
3. H. Steinhaus, Mathematical Snapshots, Oxford University Press, New York, 1969, p. 150.
4. G. Polya, Approximations to the area of the ellipsoid, Publ. Inst. Mat. Rosario, 5 (1943) 1–13.
5. D. H. Lehner, Approximations to the area of an n-dimensional ellipsoid, Canad. J. Math., 2 (1950) 267–282.
6. B. C. Carlson, Some inequalities for hypergeometric functions, Proc. Amer. Math. Soc., 17 (1966) 32–39.

Open Conference, conducted immediately after the N.S.F. funded Invitational Conference in March 1970, so that most of the CRICISAM instructor participants could remain and join others in considering new horizons for students who have completed a computer-oriented calculus sequence.

(1) CRICISAM (Center for Research in College Instruction of Science and Mathematics) Calculus: *A Computer-Oriented Presentation, Parts 1–5,* Florida State University, 1968–69. [The 1970 Edition has been combined into 2 volumes.]

(2) Committee on the Undergraduate Program in Mathematics (CUPM), Newsletter: Calculus with Computers, August 1969.

(3) Committee on the Undergraduate Program in Mathematics (CUPM), *A General Curriculum in Mathematics for Colleges* (GCMC) 1965.

(4) E. P. Miles, Jr., A Summary of the ACM-FSU *Symposia on the Impact of Computing on Undergraduate Mathematics Instruction* CACM, May 1966, pp. 388–389.

(5) E. P. Miles, Jr., et al. The papers of Atchison, Givens, Macon and Murray from (4) plus a detailed summary of the other papers and the discussion at the Symposia. CACM, September 1966, pp. 662–670.

(6) A. Ralston, A critical review of the papers in (5). Computing Reviews, November-December 1966, pp. 462–463.

(7) Proceedings of an Invitational Conference on Calculus and the Computer, CRICISAM and Florida State University, 1970, mimeographed, 36 pages.

A SURVIVAL KIT FOR THE COLLEGE MATHEMATICIAN

HARLEY FLANDERS, Purdue University

1. The problem. I believe that those who teach mathematics have a professional obligation to stay alive as mathematicians. The man teaching in a college, away from the frontier of research, sometimes does not clearly understand this obligation, but more often is under various pressures to do nothing about it.

For contrast, let us first look at the mathematics department in a research oriented university. The professor, under constant publish-or-perish pressure, assigns first priority to doing research. He may not devote sufficient time to class preparation (except for courses and seminars likely to produce research students). He doesn't give a hoot about administration.

Because his effectiveness in research is so dependent on keeping in touch, he spends much of his time in mathematical bull sessions, scanning new journals in the library, getting colloquium speakers, participating in seminars, and attending Society meetings.

In this atmosphere, the assistant professor does no administration; the associate professor on the make in research does almost no administration; relatively few full professors—and associate professors who have abandoned the research rat race—handle almost all of the administrative chores, internal and external to the department. (It is often surprising how much those who represent the mathematics department on university-wide committees and legislative bodies are *not* representative of the department—the research people won't take those assignments.)

In sum, the research man will probably assign these priorities to his working

time: (1) doing research, (2) reading the literature, bull sessions, seminars, meetings, etc., (3) graduate teaching, (4) undergraduate teaching, (5) administration.

Now let us look at the college mathematics department. The professor there gives, and certainly should give, first priority on his time and energy to (undergraduate) teaching. What I am afraid happens too often is that all his remaining time and energy gets channeled into administration.

I consider this wrong, and believe that staying alive as a mathematician and nurturing a vital interest in mathematics should receive second priority, and be considered almost as important as teaching.

The research mathematician is loyal first to his profession and second to his employer. I expect the college mathematician to be loyal first to his college and second to the mathematics profession. But this second loyalty should be a close second, and he must not allow his college loyalty to overwhelm his responsibility to mathematics, because that will harm his most important work, teaching mathematics.

2. **Research publication.** I suspect that most college mathematics professors do not produce publishable research. The atmosphere at most colleges, the strong emphasis on teaching as primary purpose, the lack of library facilities, and so on, make it pretty hard to stay near the rapidly expanding research frontier.

A few college mathematicians manage to produce publishable research in spite of these obstacles; this is certainly commendable. Some college mathematicians put considerable energy into unpublishable research. Indeed, the MAA is under some pressure to provide an outlet for what might be called *minor research* (research not acceptable to the regular research journals because it is too elementary, too small a contribution, or too uninteresting). One of the arguments pro is that college men in experimental sciences can publish occasionally, so the college mathematician is at a disadvantage re salary and promotion. Another argument pro is that getting results even in an insignificant area is of real value to the mathematics teacher. I disagree with this second argument; I have reached the conclusion, based on personal examination of hundreds of minor research manuscripts, that much of this work is so far from what mathematics is really about, that the writers are deluding themselves and, indirectly, harming their teaching.

Much minor research consists either of axiom systems for unnatural structures, or minute generalizations of known theorems and their proofs. I believe there are far more profitable activities for the college mathematician, and shall list some later. Incidentally, Prof. Klee started the Research Problems section of the Monthly to provide a source of meaningful (dare I say relevant?) open problems. I expect that few college mathematicians will ever crack one of these problems, but I know it is far healthier mathematically to work on a concrete problem than on an abstract generalization of a generalization. . . .

3. Administration. Over the years I have visited many colleges. I am always surprised at how much administration is done by mathematicians, at how much time they spend in staff and committee meetings, and at how seriously they take their administrative problems, often to the point where there seems to be little else they talk about. Obviously this is counter-productive to mathematics and must be resisted.

Unfortunately there are both faculty members and professional administrators who see administration *per se* as the most important activity of the college. There is a strong underlying belief that you can actually legislate quality, that by sufficient work on permuting course numbers, changing prerequisites, curriculum, entrance requirements, by administrative restructuring, and by the rest of the lot, the same teachers will make infinitely more out of the same students. Nonsense!

The good university with 2000 faculty members has perhaps 200 faculty administrative activists, but the college with 150 faculty members can have such overwhelming problems that it may require 149 faculty administrative activists. I have visited small colleges where the faculty as a whole meets almost weekly, and the meetings, often at night, are 3 or 4 hours long in order to get in all of the committee reports.

We have all seen how thoughtlessly a committee is appointed. Here is a typical example: The chemistry department decides its freshmen need computer coding, so it wants the history breadth requirement dropped. The committee on committees, realizing this is a matter of basic educational philosophy, appoints a 15-man committee, representing every department, to meet 4 hours per week for 10 weeks, and then to report to a special faculty meeting. Thus $15\times 4\times 10 = 600$ man-hours are spent on the committee alone. Of course the faculty meeting (150 men$\times$4 hours$=$600 man-hours) will reject the report, the committee will reconvene, etc.

If your college allocates too many of your hours to administration, then you must educate your administrators (whether autocratic central or autonomous faculty) to the plain fact that you have other work to do besides meeting classes, preps, and committees, that the regular time parcels scheduled for your own mathematical enrichment are just as inviolable as your classroom hours, and that you are hired to be a mathematics professor, not a college administrator.

This is where your chairman comes in. I believe one of the most important functions of a mathematics department chairman is protecting his staff from administrative chores, both within and without the department. Particularly, if a staff member volunteers for excessive administrative work, the chairman must protect the man from himself. If your chairman does not see this as an important part of his job, perhaps it is time for a change.

Actually it is often wise to rotate the chairmanship every 5 years or oftener. The chairman who considers his post as permanent, sooner or later must identify his interests with the administration's interests. Instead of representing the

mathematics department and fighting for its interests, he is likely to consider himself the administration spokesman in the department.

Two final suggestions: Make a fixed rule: no college business after 5 P.M. or on weekends. If there really is an overwhelming problem facing your college (department), find another college (mathematics department) you consider better than yours, which has faced the same problem, and use its solution. Corollary: adopt Harvard's curriculum and save time.

4. Time allocation. Suppose you teach 12 contact hours per week. Your preparations, office hours, and grading will take another 12 hours, so 24 hours are committed to your primary job, teaching. Assuming you want to work a 40 hour week—no more, no less—and that you are not the chairman, what happens to the other 16 hours?

I suggest an absolute maximum of 4 hours per week per man for administration, both in and out of the department. Thus, for example, if the chairman of an 8 man department (excluding himself) needs about 8 hours per week help in the department, then he should allow a maximum of 24 hours per week from mathematics for college-wide business. He should closely scrutinize the assignment of this time, and scream if it is frittered away.

This leaves a solid 12 hour per week period for mathematical survival. The point of my article so far has been to buy this time. Now I want to spend it.

5. What to do. I take it as axiomatic that if you cease to learn mathematics and cease to work at mathematical problems, then you will lose your enthusiasm for mathematics and become first a dull and then a lousy teacher. However, if mathematics is vital to you, if a beautiful theorem and proof excites you, if solving a problem thrills you, if seeing a problem challenges you, if mathematics keeps popping into your head, then you have passed survival and you might be a great teacher. The goal then is clear. I shall now list ways to spend the 12 hours per week (I so carefully extracted) working towards the goal.

(1) Work on MONTHLY problems; try to solve some and try to propose some. This is the closest you may come to original research, and when you actually solve a problem, you get much of the same satisfaction. Other problem sources are the Mathematics Magazine and the Siam Review.

(2) Read the MONTHLY. I know (naturally) of no other source of so much useful material for the college mathematician. I suggest you read all of the Notes in each issue. Each contains a nice idea, or a little gem, and each can be worked through in two hours, often less. You will broaden considerably by reading short Notes in other than your major field.

Read one major expository article in each issue (individually, or as a department project).

(3) Have a department seminar. (Nearby colleges can cooperate.) This should meet twice a week for two hours, with a break. The first hour or hour-and-a-half should be a lecture, the rest discussion. This should be regularly scheduled, say every Tuesday and Thursday, 3–5, and should never be cancelled

for college business. The lecturing should rotate among your whole staff, and everyone should prepare in advance by reading the material to be covered. (Of course, invite your better students.)

One possible topic is current MONTHLY main articles. I think working through (part of) a book is a better idea. You must go slowly and very carefully, so that everyone keeps up. Two or three pages a meeting is good progress. Pick a book with exercises and do them. Pick concrete mathematics rather than abstract mathematics. Pick a book that is recent, but in a field in which none of your staff is an expert.

These seminars should be lively. The speaker should be interrupted whenever possible, everyone should be ready to argue. No sissy stuff!

Here is a brief list of possible books:

G. Birkhoff and G. -C. Rota, *Ordinary Differential Equations*, Ginn & Co., 1962.

P. M. Cohn, *Universal Algebra*, Harper & Row, 1965.

S. Halgason, *Differential Geometry and Symmetric Spaces*, Academic Press, 1962.

L. Hörmander, *Introduction to Complex Analysis in Several Variables*, Van Nostrand, 1966.

N. Jacobson, *Lie Algebras*, Interscience, 1962.

I. Kaplansky, *Commutative Rings*, Allyn & Bacon, 1970.

W. Magnus, A. Karras, D. Solitar, *Combinatorial Group Theory*, Interscience, 1966.

H. Pollard, *Mathematical Introduction to Celestial Mechanics*, Prentice-Hall, 1966.

E. H. Spanier, *Algebraic Topology*, McGraw-Hill, 1966.

J. Todd, *Introduction to the Constructive Theory of Functions*, Academic Press, 1963.

(4) Attend local MAA and AMS meetings. Take notes on the hour addresses and discuss them in your seminar.

(5) Have a colloquium speaker once or twice a semester. Get someone from the nearest research department and pay his expenses. (You should try to offer a \$50 honorarium, although it isn't essential—it is enough for a small college. The going rate at universities is \$75–\$100.)

Keep the visitor on campus for a day and talk *mathematics* with him. This is your chance to get help with difficult textbook problems, to get things you are having trouble with explained, to hear some new ideas on classroom presentations, and so on. The chances are that your visitor would rather talk mathematics than anything else, and that he wants to help you.

(6) If there is a university less than an hour's drive from you, it might be worthwhile going as a group to a weekly seminar. If it is more than two hours away, your total time (5 or more hours) is not worth a single hour. Again, if the university is close by, you should attend the colloquium lectures when you *know* the invited speaker is a good expositor. Warning: The young colloquium speakers on interviews tend to snow everybody with their deep, deep research results.

(7) Keep abreast of professional activities by reading the Mathematical

Education section and national meeting reports in the MONTHLY, the CUPM reports, the news items and letters in the AMS Notices, and the CBM's Newsletter. This is material about mathematics, not mathematics, and is not a substitute for primary sources.

(8) Train a team for the Putnam competition. This means knowing how to work the problems yourself, and teaching your best students how to solve problems, a rewarding activity.

6. Sabbaticals. It is very important to get away from your college once in a while and devote full time to mathematics. Ideal is a sabbatical leave with full pay every seventh year. Full pay is important because college salaries are not so high that one can live for a year on half or two-thirds. And it is important to do no teaching or other paid work during that year off. Perhaps you should settle for every tenth year, but insist on full pay.

The best thing to do with your sabbatical leave is go to a major mathematical center and be a graduate student again. Audit 3 or 4 courses and seminars, and work harder than the registered students. Pick up as much mathematics as you can and hope it will keep you going for the next 10 years.

7. Final remarks. Suppose over a period of a year or two, you and your colleagues work through a hard book and work the problems. This is a fine accomplishment, and you should be proud of yourselves. No agency will hand you a monetary reward or even a brass medal. Still you should report to your college administration exactly what you are doing and why. Perhaps you will set a good example for other departments.

The goal, briefly, is to have a vital attitude toward mathematics. Do not get distracted into thinking that you are learning new things that are *directly* applicable to your undergraduate teaching. Because you are studying operators in Hilbert space does *not* mean you teach your junior first course in linear algebra as a special case of Hilbert space, nor because you are studying the Lebesgue integral do you inflict measure and content on your freshman calculus course. Restraint is the hardest thing to learn in teaching mathematics, and we are all sometimes guilty of going too far too soon. The frustration of teaching students who are years away from the things we are dying to tell them is something we have to live with, like it or not.

PROBLEMS AND SOLUTIONS

EDITED BY EMORY P. STARKE

All problems (both elementary and advanced) proposed for inclusion in this Department should be sent to E. P. Starke, 1000 Kensington Ave., Plainfield, NJ 07060. Proposers of problems are urged to enclose any solutions or information that will assist the editors. Ordinarily, problems in well-known textbooks and results in generally accessible sources are not appropriate for this Department. No solutions (except those accompanying proposals) should be sent to Professor Starke.

ELEMENTARY PROBLEMS

Solutions of Elementary Problems should be sent to Problems Group, Mathematics Department, University of Maine, Orono, ME 04473. To facilitate their consideration, solutions of Elementary Problems in this issue should be typed (with double spacing) and should be mailed before June 30, 1971. Contributors (in the United States) who desire acknowledgment of receipt of their solutions are asked to enclose self-addressed stamped postcards.

An asterisk (*) means neither the proposer nor the editors supplied a solution.

E 2283.* *Proposed by Irving Adler, North Bennington, Vermont*

Composers using the twelve-tone scale have found that for any partition of the scale into two six-tone sets A and B, the musical intervals separating pairs of tones in B are the same as the musical intervals separating pairs of tones in A, and each interval has the same multiplicity in both sets. Consider the set of integers modulo $2n$ ($Z/2n$). Partition this set into two sets A and B of n integers each. Show that the set of all differences including multiplicity (taken mod $2n$) is the same in each set.

E 2284. *Proposed by A. W. Walker, Toronto, Canada*

If a, b, c are positive numbers and $x=(b+c-a)$, $y=(c+a-b)$, $z=(a+b-c)$, show that $abc\sum yz \geqq xyz\sum bc$. Is $abc\sum bc\geqq xyz\sum yz$?

E 2285. *Proposed by A. W. Walker, Toronto, Canada*

If X, Y, Z are similarly situated points of directly similar coplanar triangles DCB, CEA, BAF annexed to any triangle ABC, then triangle XYZ is directly similar to the annexed triangles.

E 2286.* *Proposed by E. T. H. Wang, University of British Columbia*

For each positive integer n, define $f(n)$ as $f(n)=(n!)^{1/n}$. Prove or disprove that the sequence

$$\left\{\frac{f(n+1)}{f(n)}\right\}_{n=1}^{\infty}$$

is monotonically decreasing.

E 2287. *Proposed by Erwin Just and Norman Schaumberger, Bronx Community College*

If P is a nonconstant polynomial with integral coefficients and k is any integer, must there exist an integer m for which there are at least k distinct prime divisors of $P(m)$?

E 2288. *Proposed by John Corcoran, State University of New York at Buffalo*

Let L be the set of sentences of any predicate logic whose logical symbols are: the universal and existential quantifiers, identity, negation, conjunction, disjunction, implication. Does every inconsistent set of sentences from L contain at least one negative sign?

SOLUTIONS OF ELEMENTARY PROBLEMS

Nesting Habits of the Laddered Parenthesis

E 1903 [1966, 666; 1970, 525]. *Proposed by George Eldredge, El Cerrito, California*

Let an n-ladder of twos, L_n, be defined as follows:

$$L_n = 2^{2^{2^{\cdot^{\cdot^{\cdot^{2}}}}}}$$

where there are n twos. Let N_n be the number of distinct integers that can be obtained from L_n by the appropriate insertion of a set of unambiguous nested parentheses. For example, $N_3=1$, $N_4=2$. Find N_n.

Comment by E. F. Schmeichel, Itasca, Illinois. The published solution is incomplete, and in fact may be incorrect. The situation is as follows:

Suppose that $L_{1,n}, L_{2,n}, \cdots, L_{N_n,n}$ are the distinct values of n-ladders. Then the admissible $(n+1)$-ladders are of the form $(L_{i,n})^2$ or $2^{(L_{i,n})}$ for $i=1, 2, \cdots, N_n$. Certainly $2^{(L_{i,n})}>(L_{i,n})^2$ for each i; but it is not true that $2^{(L_{i,n})}>(L_{j,n})^2$ for all $i, j: i\neq j$. Hence it is conceivable that for some $i, j\in\{1, 2, \cdots, N_n\}$, $2^{(L_{i,n})}=(L_{j,n})^2$.

The current status of the problem then is:

(a) Given $L_{1,n}, L_{2,n}, \cdots, L_{N_n,n}$, distinct n-ladders of twos, prove or disprove: $2^{(L_{i,n})}\neq(L_{j,n})^2$ for all $i\neq j$, $n>5$.

(b) If the statement in (a) is true, the previous solution, $N_n=2^{(n-3)}$ is true. Otherwise the original question remains to be resolved.

Special Orthogonal Latin Squares

E 2228 [1970, 402]. *Proposed by C. C. Lindner, Emory University*

Call a latin square of order n an X_n-latin square, provided that each of the n symbols on which it is based occurs on each of the two diagonals. Show that if $n=2^k$, $k\geqq 2$, there exist $n-2$ mutually orthogonal X_n-latin squares.

Solution by D. M. Brown, Student, Harbor Beach (Michigan) Community School. The usual method for the construction of mutually orthogonal latin squares using Galois fields of order 2^k can be applied here. If A is an element of $GF(2^k)$, with $A\neq 0$ or 1, then the operation table of $x=Ay$ is an X_n-latin square. The fact that it is a latin square is well known. To prove it is an X_n-latin square, we see that any element C of $GF(2^k)$ will appear on the main diagonal when $x=y=C(A+1)^{-1}$, and on the other diagonal when $x=m-y=(C-Am)(1-A)^{-1}$, where the elements are numbered from 0 to $m=n-1$.

The fact that any two latin squares constructed in this manner are orthogonal is well known (Beck, Bleicher and Crowe, *Excursions into Mathematics*, p. 278). Since there are n elements overall and two elements do not generate X_n-latin squares when substituted for A, there must be a total of $n-2$ mutually orthogonal X_n-latin squares left.

Also solved by Tom Brylawski and Mike Vitale, A. Hedayat, David Kelly, E. F. Schmeichel, and the proposer.

Generalization of Least Common Multiple of Several Integers

E 2229 [1970, 402]. *Proposed by J. V. Michalowicz, The Catholic University of America*

Any two positive integers a, b have a g.c.d. (a, b) and an l.c.m. $[a, b]=ab/(a, b)$. Can this be generalized to express the l.c.m. of any finite number of elements in terms of g.c.d.'s only?

I. *Solution by D. C. B. Marsh, Colorado School of Mines.* For a set of n positive integers, let P_j denote the product of the g.c.d.'s of each of the $C(n, j)$ subsets of j of these integers. The l.c.m. is then given by $(P_1P_3P_5\cdots)/(P_2P_4P_6\cdots)$. E.g., for three integers,

$$[x, y, z]=\frac{x\cdot y\cdot z\cdot(x, y, z)}{(x, y)\cdot(y, z)\cdot(z, x)}\cdot$$

For a proof, consider any prime p whether or not it occurs vacuously in some or all of the factorizations of the n integers. Let the nonnegative integral powers to which it occurs be labelled $e_n\leqq e_{n-1}\leqq\cdots\leqq e_1$. Then, in P_j the prime p will have e_t $(t=1, 2, \cdots n)$ as its exponent in $C(t-1, j-1)$ cases. Thus the total exponent of p in $\prod P_{(\text{odd})}/\prod P_{(\text{even})}$ will be e_1 since the alternating sum of the binomial coefficients is zero with the exception of $C(0, 0)=1$. The stated formula follows from the multiplicative nature of the g.c.d. and l.c.m.

II. *Solution by G. A. Edgar, Student, University of California, Santa Barbara.* In any commutative lattice-ordered group [see Garrett Birkhoff, *Lattice Theory.* AMS Colloquium volume 25, chapter 13], we have

$$\bigvee_{i=1}^{n} a_i = \left[\bigwedge_{i=1}^{n} a_i^{-1}\right]^{-1} = \left[\prod_{j=1}^{n} a_j\right]\left[\bigwedge_{i=1}^{n} \prod_{j\neq i} a_j\right]^{-1}. \tag{1}$$

The positive rationals with ordinary multiplication and the partial order defined by

$$r \leqq s \quad \text{if and only if } s/r \text{ is an integer}$$

form a commutative lattice-ordered group in which for integers $x_1, \cdots, x_n$, we have $\bigvee_{i=1}^{m} x_i = \text{l.c.m.}\{x_1, \cdots, x_m\}$. $\bigwedge_{i=1}^{m} x_i = \text{g.c.d.}\{x_1, \cdots, x_m\}$. [*ibid.*, example 6, p. 293]. Hence (1) solves the problem. For example:

$$[a, b, c] = \frac{abc}{(ab, bc, ca)}; \qquad [a, b, c, d] = \frac{abcd}{(abc, abd, acd, bcd)}.$$

Note. Other solutions can be obtained from (1) by use of lattice-ordered group identities; e.g. $[a, b, c] =$

$$\frac{abc}{(a(b, c), bc)} = \frac{abc}{(a(b, c), b(c, a), c(a, b))} = \frac{abc(a, b, c)}{(a, b)(b, c)(c, a)}.$$

Also solved by D. W. Ballew, Merrill Barnebey, M. T. Bird, D. M. Bloom, Jordi Dou (Spain), R. Garfield, Ray Glenn, Michael Goldberg, Bob Gray, M. G. Greening (Australia), F. T. Howard, R. S. Matulis, Bob Prielipp, Henry Ricardo, Kenneth Rosen, George Schillinger, E. F. Schmeichel, Ralph Schreiber, N. T. Sheth, R. Sivaramakrishnan (India), Paul Smith, Stephen Spindler, A. M. Vaidya & V. S. Joshi (India), and the proposer.

Note. Many different variations were received. Ricardo observes that Uspensky & Heaslet, *Elementary Number Theory*, presents an inductive process for finding $[a_1, \cdots, a_n]$. He also finds a result equivalent to Solution I above in *Revista Math. Hispano-Americana* (4) 25(1965) pp. 235–237. Sivaramakrishnan refers to a note, *A generalization of the relation* $[m, n]\ (m, n) = mn$, The Mathematics Student, 37(1969) 194–195.

An Urn Problem

E 2230 [1970, 402]. *Proposed by B. E. Rodden, Defence Research Board, Toronto, Canada*

(A) Urns I and II each contain exactly n balls numbered consecutively from 1 to n. One ball is drawn from urn I. Balls are then drawn one at a time from urn II until the same-numbered ball is found. No balls are replaced in the urns. The process is repeated until urn II is empty. Find p_{rn} the probability of precisely r matches.

(B) For a more complex problem, m balls are drawn from urn I. Again balls are drawn one at a time from urn II and compared with the balls from urn I. When a match is made, one new ball is drawn from urn I to replace the matched ball, and the process is continued. When urn I is empty, the process continues with the remaining unmatched balls until urn II is empty. Find p_{rn} the prob-

ability of precisely r matches in the case $m=2$.

Solution by Peter Gottlieb and Harry Lass, California Institute of Technology. Without loss in generality assume that the balls are drawn from urn I in their natural order. Let N_{rn} denote the number of ways in which at least r matches occur.

$$\text{(A)} \qquad N_{rn} = \binom{n}{r}(n-r)! = \frac{n!}{r!},$$

since we need only choose r of n positions to place the integers $1, 2, \cdots, r$ in their natural order, while the remaining integers can be permuted in $(n-r)!$ ways. Thus

$$p_{rn} = \frac{1}{r!} - \frac{1}{(r+1)!} = \frac{r}{(r+1)!}, \qquad r < n,$$

with $p_{nn} = 1/n!$.

(B) For the case $m=2$, it is clear that $N_{nn}=2^{n-1}$ since there are two choices at every step of the way except for the last choice. We obtain $N_{n-1,n}$ as follows: If either of the integers 1 or 2 occurs on the first draw, we must subsequently obtain at least $n-2$ matches in the remaining $n-1$ draws. If the integer $k \geqq 3$ occurs on the first draw we will have two choices to obtain a match on each draw from the following $k-2$ draws. After that, only one choice per draw is left. Thus

$$N_{n-1,n} = 2N_{n-2,n-1} + \sum_{k=3}^{n} 2^{k-2} = 2N_{n-2,n-1} + 2^{n-1} - 2.$$

Since $N_{2,3}=6$, iteration of the above gives $N_{n-1,n}=(n-2)2^{n-1}+2$.

For $r<n-1$ we note that $N_{rn}=nN_{r,n-1}$ since ball number n can be placed in any of the n positions without contributing a match. Thus

$$N_{rn} = n(n-1)\cdots(r+2)N_{r,r+1} = \frac{n!}{(r+1)!}\left[(r-1)2^r + 2\right].$$

It follows that $p_{nn}=2^{n-1}/n!$,

$$p_{n-1,n} = \frac{1}{n!}\left[2 + (n-3)2^{n-1}\right], \quad \text{and}$$

$$p_{rn} = \frac{r+1}{(r+2)!}\left[2 + (r-2)2^r\right], \qquad r < n-1.$$

Also solved by E. F. Schmeichel, and by the proposer.

Three Coincident Centroids in a Pentagon

E 2231 [1970, 403]. *Proposed by M. S. Klamkin, Ford Scientific Laboratory*

It is a known result that if the centroid of the vertices and the centroid of the area (both uniformly weighted) of a quadrilateral coincide, then the figure is a

parallelogram. If the centroids of the vertices, of the edges, and of the area (all uniformly weighted) of a pentagon all coincide, must the figure be a regular pentagon?

Solution by W. G. Wild, Wisconsin State University. The answer is no.

Consider the pentagon with vertices at $(\pm 7k/20, 0)$, $(\pm k/2, 4/7)$ and $(0, 1)$. The centroids of the area and of the vertices coincide at $(0, 3/7)$. The centroid of the edges is located at the solution of the equation

$$\frac{3}{7}=\frac{\frac{11}{7}\sqrt{\left(\frac{k}{2}\right)^2+\left(\frac{3}{7}\right)^2}+\frac{4}{7}\sqrt{\left(\frac{3k}{20}\right)^2+\left(\frac{4}{7}\right)^2}}{\frac{7}{10}k+2\sqrt{\left(\frac{k}{2}\right)^2+\left(\frac{3}{7}\right)^2}+2\sqrt{\left(\frac{3k}{20}\right)^2+\left(\frac{4}{7}\right)^2}}.$$

(The first moment of the edge masses about the x-axis divided by the total edge mass.) These turn out to be $k=\pm 1.04228$ and ± 2.59575.

A more general solution is provided by studying the pentagon with vertices at $(\pm ak/2, 0)$, $(\pm k/2, b)$, $(0, 1)$. The centroid of the vertices is at $(0, (2b+1)/5)$ and if a is equal to $(2-b)/(3b+b^2)$, then the area centroid coincides with that of the vertices, designated by $(0, \bar{y})$, and the relation (analogous to the one in the special case above) which expresses $\bar{y}$ as a funtion of k assures us that the centroid of the edges can be made to coincide if the equation has solutions. The related existence study is routine but tedious.

Also solved by Don Coppersmith, Huseyin Demir (Turkey), and Harry Lass.

Coppersmith cites the pentagon with vertices $(0,2\sqrt{19})$, $(\pm 9, \sqrt{19})$, and $(\pm 4, -2\sqrt{19})$, along with a more general form. Lass gives $(\pm a, 0)$, $(\pm a, 1)$ and $(0, 1+\frac{1}{3}\sqrt{6})$, where a is the positive root of a certain quadratic equation.

Representation of 1 by Egyptian Fractions

E 2232 [1970, 403]. *Proposed by H. D. Ruderman, Hunter College High School, New York City*

Let U_n be the smallest number of different unit fractions totalling 1 where the largest unit fraction is $\leqq 1/n$. For example, $U_1=1$, $U_2=3$, and $U_3=5$ because

$$1=\frac{1}{3}+\frac{1}{4}+\frac{1}{5}+\frac{1}{6}+\frac{1}{20}\cdot$$

It is well known that for every n there is a finite number of terms giving the sum 1. Find an upper bound on U_n.

Solution by P. Erdös and E. Straus. There are constants c_1 and c_2 so that

$$(1)\qquad (e-1)n-c_2<U_n<(e-1)n+c_1n/\log n.$$

The first inequality immediately follows from

$$\sum_{t=a}^{b} \frac{1}{t} < \log b - \log a + \frac{c}{a} \cdot$$

To prove the second inequality define m by

$$(2) \quad \frac{1}{n} + \frac{1}{n+1} + \cdots + \frac{1}{m} < 1 < \frac{1}{n} + \frac{1}{n+1} + \cdots + \frac{1}{m} + \frac{1}{m+1} \cdot$$

Clearly $m = en + O(1)$. Put

$$(3) \quad \frac{u}{v} = 1 - \frac{1}{n} - \frac{1}{n+1} - \cdots - \frac{1}{m}, \qquad 0 < \frac{u}{v} < \frac{1}{m+1} \cdot$$

Clearly v is less than or equal to the least common multiple of the integers not exceeding m, so that

$$(4) \qquad v < m^{\pi(m)} < e^{2m}.$$

Now, a theorem of Erdös (*On the solutions in integers of*

$$\frac{a}{b} = \frac{1}{x_1} + \cdots + \frac{1}{x_n},$$

in Hungarian, Mat. Lapok 1(1950), 192–210) states that for every $1 \leqq u < v$,

$$(5) \quad \frac{u}{v} = \frac{1}{x_1} + \cdots + \frac{1}{x_k}, \qquad x_1 < \cdots < x_k, \quad k < c \log v / \log\log v,$$

is always solvable in integers $x_1, \cdots, x_k$. From (2), (3) and (5), $x_1 > m+1$; thus

$$1 = \frac{1}{n} + \cdots + \frac{1}{m} + \frac{1}{x_1} + \cdots + \frac{1}{x_k}, \qquad m - n + k < (e-1)n + c_1 n/\log n,$$

which completes the proof of (1).

It seems to us certain that $U_n - (e-1)n \to \infty$ as $n \to \infty$ but we have not proved this.

Bounds were also found by Haig Bohigian, C. Gardner, David Kelly, O. P. Lossers (Netherlands), Simeon Reich (Israel), E. F. Schmeichel, and the proposer. Several solvers conjecture that $2n$ is an upper bound for U_n.

A related paper (communicated to us by the late Leo Moser) is H. Saltzer, this MONTHLY, (1947) 135–142. Reich calls attention to J. C. Owings, Jr., *Another proof of the Egyptian Fraction theorem*, this MONTHLY, (1968) 777–778.

Points Minimizing the Sum of Geodesics

E 2233 [1970, 403]. *Proposed by E. J. Cockayne, University of Victoria, Canada*

Let A, B, C be three distinct points on the surface of a sphere, not all on the same great circle. The closed curve Γ formed by the minor great circle arcs

AB, BC, CA divides the surface into two unequal areas. Suppose Z is the set of points which compiise the smaller area including the boundary Γ. Prove that any point P on the surface minimizing the sum of the minor great circle arcs $PA+PB+PC$ is a point of Z.

Solution by M. G. Greening, University of New South Wales, Australia. Throughout, AB, AP, etc. are minor arcs of great circles of lengths $|AB|$, $|AP|$, etc.

There are two possibilities for P exterior to Γ.

(1) None of AP, BP, CP intersect Γ again. Then the sum of the surface angles at A, B, C, P is 8π which equals the sum of the angles contained by the four spherical triangles APB, APC, BPC, ABC. This, however, means that at least one triangle has angle sum $\leqq 2\pi$; impossible.

(2) AP, say, intersects BC at D. Then the geodetic property gives $|BP| + |PC| > |BC|$ so that

$$|AP| + |BP| + |CP| > |AD| + |BD| + |CD|.$$

Also solved by Michael Goldberg, David Kelly, E. F. Schmeichel, and the proposer.

ADVANCED PROBLEMS

All solutions of Advanced Problems should be sent to J. Barlaz, Rutgers—The State University, New Brunswick, N.J. 08903. Solutions of Advanced Problems in this issue should be typed (with double spacing) on separate, signed sheets and should be mailed before June 30, 1971. Contributors (in the United States) who desire acknowledgment of receipt of their solutions are asked to enclose self-addressed stamped postcards.

*An asterisk (*) means neither the proposer nor the editors supplied a solution.*

5783.* *Proposed by D. A. Moran, Michigan State University*

Let Z denote the set of integers, and let p be a fixed prime. For each positive integer a, define

$$U_a(n) = \{n + \lambda p^a : \lambda \in Z\}.$$

Then, as is well known, $\{U_a(n)\}$ is a basis for some topology $\mathfrak{J}_p$ on Z. If $p \neq q$, it is easy to show that $\mathfrak{J}_p$ and $\mathfrak{J}_q$ are distinct topologies on Z, and that $(Z, +, \mathfrak{J}_p)$ and $(Z, +, \mathfrak{J}_q)$ are not isomorphic topological groups.

Prove or disprove: $(Z, \mathfrak{J}_p)$ and $(Z, \mathfrak{J}_q)$ are never homeomorphic topological spaces.

5784. *Proposed by ANON, Erewhon-upon-Wabash.*

Let $x_0 = a$, $x_1 = a+h$, $x_2 = a+2h$, $x_3 = a+3h$. Prove the existence of unique polynomials $u(x)$, $v(x)$, $w(x)$ of degree 5 such that

$$\int_{x_0}^{x_1} u(x)f^{(5)}(x)dx + \int_{x_1}^{x_2} v(x)f^{(5)}(x)dx + \int_{x_2}^{x_3} w(x)f^{(5)}(x)dx = 44\int_{x_0}^{x_1} f(x)dx$$
$$+ 155\int_{x_1}^{x_2} f(x)dx + 44\int_{x_2}^{x_3} f(x)dx$$

for each $C^{(5)}$ function f which vanishes at x_0, x_1, x_2, x_3.

5785.* *Proposed by V. A. McAuley, Marshall Space Flight Center, Huntsville, Alabama*

Show that for each choice of the natural number m there are m positive numbers $d_j (j=1, \cdots, m)$ with each $d_j > 1$, such that

$$(x+2)^{2m} + x^{2m} \equiv 2\prod_{j=1}^{m} (x^2 + 2x + d_j)$$

is an identity.

5786.* *Proposed by Jan Mycielski, University of California, Berkeley*

Find a four-chromatic graph such that at each vertex four edges meet and each edge is contained in exactly one triangle. What is the minimal number of vertices of such a graph?

5787. *Proposed by J. L. Bryant, Florida State University*

Let $\{(a_i, b_i)\}$ be a finite collection of pairs of points in the plane each satisfying $|a_i - b_i| \leqq 1$ with all points distinct. Show that each a_i can be connected to each b_i by an arc whose diameter is no greater than $\sqrt{13}$, so that no two arcs intersect. (Diameter of an arc C means $\max(|x-y|$ for $x, y \in C)$.)

5788.* *Proposed by N. S. Mendelsohn, University of Manitoba*

Let G be a group with presentation $G = gp\cdot(A, B : A = (BA)^r B, B = (AB)^s A)$. Show that G is finite for all choices of the positive integers r and s, and that either G is cyclic or G has a cyclic subgroup of index 2.

SOLUTIONS OF ADVANCED PROBLEMS

Sums of Random Variables with Infinite Expectation

5716 [1970, 197]. *Proposed by Harry Kesten, Cornell University*

Let $X_1, X_2, \cdots$ be a sequence of independent identically distributed random variables and $S_n = \sum_{i=1}^{n} X_i$. Show that

$$\text{(1)} \qquad \limsup_{n\to\infty} \frac{|X_n|}{|S_{n-1}|} = \infty \qquad \text{with probability 1,}$$

whenever $E|X_i|$, the expectation of $|X_i|$, is infinite. Use (1) to derive the fol-

lowing theorem of Chow and Robbins (Proc. Nat. Acad. Sci., 47 (1961) 330–335): When $E|X_i| = \infty$, then for any sequence $\{b_n\}$ of positive numbers either

$$\liminf_{n\to\infty} \frac{|S_n|}{b_n} = 0 \qquad \text{with probability 1}$$

or

$$\limsup_{n\to\infty} \frac{|S_n|}{b_n} = \infty \qquad \text{with probability 1.}$$

Solution by the proposer. Without loss of generality we may assume that X_i can take only the values 2^k, $k=0, 1, \cdots$. Indeed, if we replace X_i by

$$Y_i = \begin{cases} 1 & \text{if } |X_i| \leqq 1 \\ 2^{k+1} & \text{if } 2^k < |X_i| \leqq 2^{k+1}, \quad k = 0, 1, \cdots, \end{cases}$$

then $(\sum_{i=1}^{n-1} Y_i)^{-1} Y_n \leqq 2|S_{n-1}|^{-1}(|X_n|+1)$. If S_n is recurrent (1) is immediate, and if S_n is transient $S_{n-1}\to\infty$; thus

$$\limsup_{n\to\infty} \left(\sum_{i=1}^{n-1} Y_i\right)^{-1} Y_n = \infty \qquad \text{with probability 1}$$

implies (1). Assume then that $P\{X_i=2^k\}=p_k$, $k=0, 1, \cdots$, $\sum p_k=1$, $\sum 2^k p_k = \infty$. Define

$$r_k = \sum_{i=k}^{\infty} p_i, \qquad \mu_k = \sum_{i=0}^{k} 2^i p_i, \qquad n_k = [2^k(\mu_k \log \mu_k)^{-1}]$$

and let E_k be the event: $E_k = \{X_i = 2^k \text{ for some } i \leqq n_{k-1}\}$; then

$$n_{k-1}p_k \geqq P\{E_k\} = 1 - (1-p_k)^{n_{k-1}} = C_1 \min(1, n_{k-1}p_k) \tag{2}$$

for a suitable $C_1>0$. By the integral comparison test (see Knopp, *Theory and Application of Infinite Series*, §39), $\sum n_{k-1}p_k = \infty$ and thus

$$\sum P\{E_k\} = \infty. \tag{3}$$

It is easy to see that

$$P\{E_{k_1} \cap E_{k_2}\} \leqq P\{E_{k_1}\}P\{E_{k_2}\}, \qquad k_1 \neq k_2. \tag{4}$$

By a suitable form of the Borel-Cantelli lemma, (3) and (4) imply

$$P\{E_k \; i.o.\} = 1 \qquad (\text{"}i.o.\text{" means "occurs infinitely often"}) \tag{5}$$

(see A. Renyi, *Wahrscheinlichkeitsrechnung*, Hilfssatz C, Ch. VII, §5). Next let F_k be the event

$$F_k = E_k \cap \left\{ \sum_{i=1}^{n_{k-1}} X_i^{(k)} > 2^k (\log \mu_{k-1})^{-1/2} \right\},$$

where

$$X_i^{(k)} = \begin{cases} X_i & \text{if } X_i \leqq 2^{k-1} \\ 0 & \text{if } X_i \geqq 2^k. \end{cases} \tag{6}$$

Then $E\{X_i^{(k)} \mid X_i^{(k)} \leqq 2^{k-1}\} = (1-r_k)^{-1}\mu_{k-1} \leqq 2\mu_{k-1}$ for $k \geqq k_0$ and this implies, for $k \geqq k_0$,

$$E\left\{\sum_{i=1}^{n_{k-1}} X_i^{(k)} \mid E_k\right\} \leqq 2n_{k-1}\mu_{k-1},$$

$$P\left\{\sum_{i=1}^{n_{k-1}} X_i^{(k)} > 2^k(\log \mu_{k-1})^{-1/2} \mid E_k\right\} \leqq 2^{-k}(\log \mu_{k-1})^{1/2} 2n_{k-1}\mu_{k-1}$$
$$\leqq (\log \mu_{k-1})^{-1/2}.$$

Thus

$$\sum_{k \geqq k_0} P\{F_k\} \leqq \sum_{k \geqq k_0} (\log \mu_{k-1})^{-1/2} P\{E_k\} \leqq \sum_{k \geqq k_0} (\log \mu_{k-1})^{-1/2} n_{k-1} p_k$$
$$\leqq \sum \frac{2^k p_k}{\mu_{k-1}(\log \mu_{k-1})^{3/2}}.$$

We shall first prove (1) using the additional assumption:

$$2^k p_k = O(\mu_{k-1}) \quad \text{or (equivalently)} \quad \mu_k = O(\mu_{k-1}). \tag{7}$$

Then (again by the integral comparison test)

$$\sum P\{F_k\} = O\left(\sum \frac{2^k p_k}{\mu_k (\log \mu_k)^{3/2}}\right) < \infty,$$

and by the Borel-Cantelli lemma $P\{F_k \ i.o.\} = 0$. In view of (5) this gives

$$P\{E_k \backslash F_k \ i.o.\} = 1.$$

Now, if E_k occurs, but not F_k, let m_k be the smallest i with $X_i \geqq 2^k$. On E_k, $m_k \leqq n_{k-1}$, whereas on $E_k \backslash F_k$

$$S_{m_k-1} = \sum_{i<m_k} X_i = \sum_{i<m_k} X_i^{(k)} \leqq \sum_{i \leqq n_{k-1}} X_i^{(k)} \leqq 2^k(\log \mu_{k-1})^{-1/2} \leqq X_{m_k}(\log \mu_{k-1})^{-1/2}.$$

Thus when $E_k \backslash F_k$ occurs $i.o.$, then

$$\limsup_{n\to\infty} \frac{X_n}{S_{n-1}} \geqq \limsup_{k\to\infty} \frac{X_{m_k}}{S_{m_k-1}} = \infty,$$

and, assuming (7), (1) is established.

If (7) fails, we can find $k_1 < k_2 < \cdots$ such that $2^{k_j} p_{k_j} \geqq j^4 \mu_{k_j-1}$. Let

$$N_j = [p_{k_j}^{-1}], \qquad E_j' = \{X_i = 2^{k_j} \text{ for some } i \leqq N_j\},$$
$$F_j' = E_j' \cap \left\{ \sum_{i=1}^{n} X_i^{(k_j)} > j^2 \mu_{k_j-1} N_j \right\}$$

with $X_i^{(k)}$ as defined in (6). Again $\sum P\{E_j'\} = \infty$, $\sum P\{F_j'\} < \infty$ and (1) follows in the same way as before.

Derivation of the Chow-Robbins theorem from (1). Assume that

$$P\left\{\limsup_{n\to\infty} \frac{|S_n|}{b_n} < \infty\right\} = a > 0.$$

One easily sees that necessarily $b_n \to \infty$ in this case, and hence by Kolmogorov's zero-one law, $a = 1$. Also

$$\limsup_{n\to\infty} \frac{\left|\sum_2^n X_i\right|}{b_n} = \limsup_{n\to\infty} \frac{|S_n|}{b_n}$$

and

$$P\left\{\limsup_{n\to\infty} \frac{|S_{n-1}|}{b_n} < \infty\right\} = P\left\{\limsup_{n\to\infty} \frac{\sum_2^n X_i}{|b_n|} < \infty\right\} = 1.$$

Lastly,

$$P\left\{\limsup_{n\to\infty} \frac{|X_{n+1}|}{b_n} < \infty\right\}$$
$$= P\left\{\limsup_{n\to\infty} \frac{|X_n|}{b_n} < \infty\right\} = P\left\{\limsup_{n\to\infty} \frac{|S_{n-1}| + |S_n|}{b_n} < \infty\right\} = 1,$$

which, together with (1), implies

$$\liminf_{n\to\infty} \frac{|S_n|}{b_n} \leqq \liminf_{n\to\infty} \frac{|S_n|}{1 + |X_{n+1}|} \cdot \limsup_{n\to\infty} \frac{1 + |X_{n+1}|}{b_n} = 0$$

with probability 1.

A Chain of Subspaces in a Hilbert Space

5717 [1970, 197]. *Proposed by W. H. Ruckle, Lehigh University*

Without using the axiom of choice (i.e., no Hamel basis) construct a continuum $\{X_r : 0 < r \leqq 1\}$ of linear subspaces of a Hilbert space H which has the following properties: (a) X_r is dense in H for each r; (b) if $r < s$, $X_r \subset X_s$ and X_s/X_r has uncountable dimension; (c) $\bigcup_r X_r \neq H$.

Solution by V. L. Klee, University of Washington. Represent Hilbert space as $L^2[-1, 1]$. For each $r \in [0, 1]$, let X_r consist of all functions $f \in L^2[-1, 1]$ such that for some $a < r$, the restriction of f to $[a, 1]$ is equivalent to a function which assumes only finitely many values. The set $\{X_r\}$ fulfills all conditions of the problem.

Also solved by D. A. Hejhal, and the proposer.

Dense Iterate Operations

5719 [1970, 198]. *Proposed by Roger Lyndon, University of Michigan*

For $k \geqq 0$, let S be the set of all numbers of the form

$$s = \sqrt{k \pm \sqrt{k \pm \cdots \pm \sqrt{k}}}$$

with arbitrary finite sequence of signs. If $k \geqq 2$, then all s in S are real. Prove: (1) if $k=2$, then S is dense in the interval $(0, +2)$; (2) if $k>2$, then S is dense in no real interval.

I. *Solution (1) by L. S. Liverpool, Imperial College, London, England.* The set S of expressions of the form

$$t = \pm \sqrt{k \pm \sqrt{k \pm \cdots \pm \sqrt{k}}}$$

is precisely the set of zeros of the polynomial $P_n(t)$ for some $n = 1, 2, 3, \cdots$, where

$$P_1(t) = P(t) = t^2 - k, \qquad P_n(t) = P_1(P_{n-1}(t)), \quad n = 2, 3, \cdots.$$

In the case $k=2$ we may put $t = 2\cos\omega$, whence $P_n(t) = 2\cos(2^n\omega)$ and the zeros of P_n are $2\cos\{(2k+1)\,2^{-n}\pi\}$, $k=0, 1, \cdots$, so that S is a dense subset of $[-2, 2]$.

II. *Solution (2) by the proposer.* Let $k>2$. It will suffice to show that S contains none of its own accumulation points. Evidently $M = \sup S$ satisfies $M = (k+M)^{1/2}$, whence $M = \frac{1}{2}(1+(1+4k)^{1/2})$. Now $k>2$ implies $M<k$, whence $m = \inf S = (k-M)^{1/2} > 0$. Therefore the greatest s in S less than $\sqrt{k}$ is less than $(k-m)^{1/2}$, and the least s in S greater than $\sqrt{k}$ is greater than $(k+m)^{1/2}$. It is easily seen that if s has a deleted neighborhood disjoint from S, then each of $(k+s)^{1/2}$ and $(k-s)^{1/2}$ has such a neighborhood. It follows by induction that each s in S has a neighborhood containing no other point of S.

Also solved by M. T. Bird, D. Ž. Djoković, and D. A. Hejhal.

Liverpool in his proof of (2) uses the ideas of normal functions and some results in P. J. Myrberg, *Iteration der reelen Polynome zweiten Grades*, II, Ann. Acad. Sci. Fennicae, A. I. Math. Nr. 268 (1959).

A Space not Normal but with all Proper Subspaces Normal

5720 [1970, 313]. *Proposed by Otto Morphy.*

Using the definition of "normal" not containing "Hausdorff," find all topo-

logical spaces X which are not normal but are such that each proper subspace is normal.

Solution by Gerald Wildenberg, Clark University. X contains a pair of disjoint nonvoid closed sets C_1, C_2 such that if O_1 and O_2 are open sets such that $O_1 \supset C_1$, $O_2 \supset C_2$, then $O_1 \cap O_2$ is nonvoid. But for all $x \in X$, $X - \{x\}$ is normal, which implies that for all $x \in X$, $\{x\} = C_1$ or C_2 or $O_1 \cap O_2$. Thus X has both at least three points and at most three points, precisely two of which are closed. Hence X is homeomorphic to $\{1, 2, 3\}$ with the topology: $\{\varnothing, \{1, 2, 3\}, \{1, 2\}, \{2, 3\}, \{2\}\}$.

Also solved by Einar Andresen (Norway), Cleveland State University Problem Solving Group, A. A. Jagers (Netherlands), F. J. Papp, Mark Yu, and the proposer.

Fixed-Point Mappings of the n-Ball

5721 [1970, 313]. *Proposed by Simeon Reich, The Technion, Haifa, Israel*

Let $f: K^n \rightarrow E^n$ be a continuous function where E^n is Euclidean n-space and $K^n = \{x \mid x \in E^n, \|x\| \leqq 1\}$ such that for every $y \in S^{n-1} = \{x \mid x \in E^n, \|x\| = 1\}$ there is no $m > 1$ with $f(y) = my$. Show that f has a fixed point.

Solution by Henry Ricardo, Yeshiva University and Manhattan College. If f has no fixed point, then the function $g: K^n \rightarrow S^{n-1}$ given by $g(x) = (f(x) - x) / \|f(x) - x\|$ is continuous.

By the Brouwer Fixed Point Theorem, there must exist a point y such that $g(y) = (f(y) - y)/\|f(y) - y\| = y$, implying that $y \in S^{n-1}$ and $f(y) = y\{1 + \|f(y) - y\|\} = my$, where $m > 1$, a contradiction.

Also solved by Einar Andresen (Norway), B. H. Aupetit, Alan Berger, Josef Daneš (Czechoslovakia), E. P. Del Norte, Crist Dixon, D. Ž. Djoković & W. J. Gilbert, Robert Fraga (Lebanon), R. V. Fuller, W. J. Gilbert, M. L. J. Hautus (Netherlands), E. C. Hook, A. A. Jagers (Netherlands), Douglas Lind, O. P. Lossers (Netherlands), Beatriz Margolis (Argentina), P. J. Owens (England), T. M. Phillips & J. C. Liggett, Otto Platt, Paul Smith, J. W. Thomas, Konrad Victor (Israel), R. M. Warten, W. T. Whitley, and the proposer.

For more general results implying the theorem of this problem, Berger and Thomas refer to W. V. Petryshyn, *On nonlinear p-compact operators in Banach spaces with applications to constructive fixed-point theorems*, Journal of Math. Anal. and Appl., 15 (1966) 228–242.

The Supremum of Monic Polynomials

5722 [1970, 313]. *Proposed by D. G. Cantor, University of California, Los Angeles*

Let X be a compact subset of the reals. Prove that the necessary and sufficient condition for the existence of a nonconstant monic polynomial with real coefficients which has absolute value < 1 on X is that there exists such a polynomial which has absolute value < 2 on X. Show that "2" is sharp; i.e., 2 cannot be replaced by any larger number.

Solution by Einar Andresen, University of Oslo, Norway. Let $P(X)$ be the set of polynomials defined on X; $P(X)$ is given the sup norm. Define a nonlinear

operator $A: P(X) \to P(X)$ by

$$A f(X) = f(x)^2 - \tfrac{1}{2}\|f\|^2, \qquad f \in P(X), \quad x \in X.$$

If f is monic, so is Af. We have $\|Af\| = \frac{1}{2}\|f\|^2$, and by induction:

$$\|A^n f\| = 2\|f\|^{2^n}/2^{2^n}.$$

If $\|f\| < 2$, it follows that $\|A^n f\| < 1$ for some n. The first assertion is proved.

To prove the second, let $X = [-2, 2]$. The identity polynomial has norm 2 on X. We show that any monic polynomial has norm at least 2 on X; thus we prove that "2" is sharp.

Suppose $f \in P(X)$ is monic, $\|f\| < 1$. To obtain a contradiction we shall show: (1) There is an operator $T: P(X) \to P(X)$ which takes monic polynomials of degree $2k$ into monic polynomials of degree k, such that $\|TP\| \leqq \|P\|$: (2) For some $n \in N$ there exists a monic polynomial g of degree 2^n with $\|g\| < 2$. We would then have a monic polynomial h on $[-2, 2]$ of degree 1, $\|h\| < 2$, which is impossible.

Proof of (1). Let $h \in P(X)$. Define $h' \in P(X)$ by $h'(x) = \frac{1}{2}(h(x) + h(-x))$, $x \in X$. h' can be considered as a polynomial in x^2. Define $T'h$ by $T'h(x^2) = h'(x)$, $x \in X$; $T'h$ is a polynomial defined on $[0, 4]$. Finally, define Th by $Th(x) = T'h(x+2)$, $x \in X$. It is clear that the operator T has the desired properties.

Proof of (2). Suppose f has degree $2^m p$, p odd. The polynomial $t = T^m f$ is monic, has degree p, and $\|t\| < 1$. There is an $n \in N$ such that p divides $2^n - 1$. Write $2^n - 1 = pq$; and define a monic polynomial $g \in P(X)$ by $g(x) = t(x)^q \cdot x$. Then $\|g\| < 2$, and the proof is complete.

Also solved by D. F. Behan, D. A. Hejhal, O. P. Lossers (Netherlands), Konrad Victor (Israel), J. E. Wilkins, Jr., and the proposer.

Several solvers obtain the solution using known properties of Tchebycheff polynomials.

REVIEWS

EDITED BY J. ARTHUR SEEBACH, JR. AND LYNN A. STEEN

with the assistance of the mathematics departments of St. Olaf and Carleton Colleges

COLLABORATING EDITOR FOR FILMS: SEYMOUR SCHUSTER, Carleton College

Printed materials for review should be sent to: Book Review Editor, American Mathematical Monthly, St. Olaf College, Northfield, MN 55057. Films and correspondence relating to films should be sent to Seymour Schuster, Carleton College, Northfield, MN 55057.

All unsigned material is written by the editors. A boldface capital **C** *in the margin indicates that a review is based in part on classroom use. Professors willing to write such a review should inform the editor in order to avoid duplication.*

Calculus. By M. Evans Munroe. Saunders, Philadelphia, 1970. ix+763 pp. $12.75. (Telegraphic review, June–July 1970.)

This three-semester beginning calculus text is distinguished by the following unusual features. It proceeds on plausibility and intuition rather than detailed

proofs. (To use the author's historical metaphor, it is eighteenth century calculus with twentieth century improvements, leaving nineteenth century scruples to a later course, where they belong.) It revives the language of variables and differentials to complement the language of functions and derivatives, the former being the natural language of most applications. (Differentials are not infinitely small increments, however, but linear forms on the tangent spaces. This means the subject is differentiable manifolds.) It is fussily precise about notations, distinguishing for instance between the equation $y=f(x)$ defining a locus and the identity $y=f(x)$ which then holds on the locus. (But both would be written $y=f\circ x$, the variables x and y being in one case mappings of $R^2\to R$, and in the other of the locus to R.) The multidimensional later part makes full use of the linear algebra which is included, and culminates in an elegant treatment of Stokes' Theorem for exterior differential forms and its specializations.

All this sounds refreshing and promising, and indeed the book deserves appreciative comments on many excellent points. In the end, however, it stands or falls on its most novel aspect, which is the explanation of differentials. This review will focus on that.

The reviewer agrees with Munroe that beginning calculus students could and should learn to use variables and differentials correctly, even if this poses a new teaching problem. While the only safe way to measure the success of an innovative book is to use it with live students, as the author has done, the reviewer must venture the estimate that this attempt is not as skillful and sensitive to student psychology as it needs to be. Improvements are certainly possible. Two specific complaints are that the author's conception of tangent space is unnecessarily far-fetched, and that the presentation is too geometric in its bias, to the exclusion of applications from other fields.

The author has previously stated his case (this MONTHLY, 65 (1958) 81–90, reprinted in *Selected Papers on Calculus*, Mathematical Association of America, 1969). After derivatives of functions have been defined and the notation $D_u v=f'\circ v$ (where $v=f\circ u$) adopted, differentials first appear in a geometric setting. The manifold here is a smooth curve C in the plane, with coordinate variables x and y and an arc length variable s on it. We are told that dx and dy at $p\in C$ are the variables on the tangent line T_p at p obtained by restricting $x-x_p$ and $y-y_p$ to T_p, and ds is length on T_p. Good. The problem is to define differentials of other variables. The goal is, of course, to make for each $u: C\to R$ a variable $du: T_p\to R$ such that the "fundamental theorem of differentials" holds, namely $dv=D_u v du$ when $v=f\circ u$. The reviewer's response to this situation (and the student's?) would be that, since this relation *is* fundamental, and since it in effect gives you dv explicitly in terms of dx or dy, or ds or any other local coordinate, it should be the definition. One would immediately use the chain rule for derivatives of functions to show that the choice of a local coordinate in the definition makes no difference.

Munroe chooses instead to make a big production of a twentieth century "breakthrough" 250 years after Leibniz, which is the discovery that the tangent line T_p is really the set of all "derivative operators" at p, namely all linear maps

D: {variables}$\to R$ which obey the Leibniz rule: $D(uv)=u_pDv+v_pDu$. Then du: $T_p\to R$ is defined by $du(D)=Du$. In doing this he strains the naive understanding in two unnecessary ways. First, having presented the manifold as a geometric object, he defines its tangent line as an algebraic object only tenuously related to the tangent line which you see. Secondly, by shifting the axiom role from the fundamental theorem to Leibniz's rule, he makes trouble of a nineteenth century sort in that now the proof of the former from the latter requires that the manifold be of class C^3, whereas the proof in the reverse direction requires no such awkward assumption. (You apply $D(w^2)=2wD_pw$ to the squares in $4uv=(u+v)^2-(u-v)^2$.) The rabbit-out-of-a-hat nature of Munroe's approach seems to the reviewer to be destructive of a true appreciation of how mathematics is created.

Teaching this subject at this level needs to make full use of all motivating resources in the student's experience. These include both geometric feeling for curves and surfaces, and physics, the solving of applied problems. Munroe does not make full use of the second; he presents the material with purely geometric motivation, and only much later shows some of its central applications. Such a bias could hamper the student in learning how naturally manifolds fit many physical situations. Beginning concretely with a curve in the plane is certainly sound, but Munroe is at pains to give a geometric genesis of the coordinate variables x and y on it, never suggesting that in practice x and y would probably already have physical meanings, so that the curve is the image of a set of physical states under a mapping into R^2 given by the pair (x, y). This is important because the same physical situation can be examined via other variables, and so seen as a different curve in a different space, a flexibility which is often needed. Note that arc length need not have physical significance.

To sum up the overall effect in a highly subjective judgment, calculus comes through as a tangle of interlocking and redundant notations, mostly not suggestive of concrete experience. Seeking something to hold onto, the student can be expected to learn some formulas and methods, and be able to solve the problems which he recognizes. The fact that this description also fits most traditional calculus courses unfortunately does not justify this text. It is much to be hoped that this effort will be used as an experimental beginning and an inspirational spur for other attempts. If so, we shall owe much to M. Evans Munroe.

F. CUNNINGHAM, JR., Bryn Mawr College

C *The Elements of Complex Analysis*. By J. Duncan. Wiley, New York, 1968. ix+313 pp. \$11.50 (cloth), \$5.75 (paper). (Telegraphic Review, March 1969).

Here is a text which presents complex analysis from an elegant and rigorous modern-analytic point of view, very much in the spirit which serious students appreciate, yet which is quite accessible to the average student. The text was used by the reviewer at Reed College for a semester-length course for senior math majors, most of whom had a good background in elementary real analysis and modern algebra.

Duncan begins with a very nice chapter on metric topology, followed by a careful introduction to the complex and extended complex planes. The standard theory of differentiation and power series come next. Now comes the heart of the book. Very much in the spirit of Rudin's *Real and Complex Analysis*, Duncan proves the Cauchy theorem for star-shaped regions. This form of the theorem avoids the Jordan curve theorem, has a beautiful and straightforward proof, and includes nearly every case for which the theorem is used. Duncan then treats the standard function theory results, most of which are a consequence of the Cauchy theorem, by dividing them into those which are of a local nature and those which are global.

If one could find fault with the text it would be with the exercise sets. Each of the problems is a challenging project, and it would be useful to have more problems of a routine type which would illustrate the theorems in the text. Some of the problems, which on the surface seem to do this, turn out to involve unnecessarily long and cumbersome computations. This reviewer found himself making up a sizable number of supplementary problems. A few of the problems are not even solvable at the level at which they are presented. For example, in the power series section (before any integration theory) the student is asked essentially to prove Liouville's theorem: "Which power series functions are differentiable at infinity?"

Overall, however, the class and I found the book completely satisfying—an elementary yet modern and rigorous introduction to complex analysis.

L. A. EDISON, Alma College, Michigan

C *Linear Algebra and Geometry. A Second Course.* By Irving Kaplansky. Allyn and Bacon, Boston, 1969. 151 pp. $9.75. (Telegraphic Review, January 1970.)

This book contains an elegant treatment of topics from linear algebra and geometry which the author has taught in a two term sequence at the University of Chicago. The first two chapters are devoted to linear algebra and the third chapter to geometry.

It is assumed that the reader has had a first course in linear algebra and is "comfortably acquainted with the elements of linear algebra, done in a coordinate-free style starting with abstract vector spaces." In addition to the usual topics on inner product spaces, the first chapter contains the Witt cancellation theorem, hyperbolic planes, quadratic forms over fields of characteristic two and forms over rings. The second chapter is entitled Orthogonal Similarity and is devoted to the study of linear transformation on an inner product space V.

As indicated in the preface, linear algebra books usually stop short of the interesting applications of linear algebra to geometry and few geometry books even recognize linear algebra. As the author states, "Classical geometry, linear algebra's twin sister, is a bridesmaid whose chance of getting near the altar becomes ever more remote." However, in Chapter 3, classical geometry has not become the bride. Affine planes and inner product planes are studied briefly,

but the real emphasis is on projective planes. A concise but elegant treatment is given of projective transformations, duality, cross ratio and harmonic range, and conics. Each topic is presented from the point of view of linear algebra.

In the judgment of the reviewer, this text is best suited for honors undergraduate students or beginning graduate students. The concise exposition, the sparseness of concrete examples and the exercises would seem to limit the use of this book as a text. Nevertheless, anyone searching for a textbook on linear algebra and geometry for capable students would do well to give this book careful consideration.

R. J. TROYER, Lake Forest College

*Studies in Geometry.** By Leonard M. Blumenthal and Karl Menger. Freeman, San Francisco, 1970. xiv+512 pp. $15.00. (Telegraphic Review, January 1971.)

This is a very elementary, meticulously formal introduction into a few chapters of geometry which in its kind can hardly be surpassed, with many useful exercises to develop by little steps reasoning in these fields. One of the authors in his preface recalls the famous words that Plato wrote on the entrance gate to his academy: "*Let no one unacquainted with geometry enter here.*" The referee would prefer here to quote another porch inscription "*Lasciate ogni speranza voi ch'entrate,*" though there is some hope left for the reader that after a longwinded path through *Inferno* and *Purgatorio* he enjoys the *Paradiso* of geometry.

HANS FREUDENTHAL, Utrecht, Netherlands

Geometry. By William R. Ballard. Saunders, Philadelphia, 1970. xii+238 pp. $9.00. (Telegraphic Review, March 1970.)

"If two sides of a triangle are congruent, then the angles opposite these sides are congruent." Once, this version of the base angles theorem would have been regarded as much too fussy and pedantic. Today, alas, with the meaning of the word "equal" temporarily stabilized, it is part of the language and, in fact, Theorem 6.1 of the present text.

The text itself is a reasonable account of elementary geometry intended for consumption by elementary education majors. The major part of the text consists of the axioms and theorems of Euclidean geometry (mostly in the plane). Birkhoff's so-called "ruler postulate" is assumed, as well as an analogous axiom for measuring angles. In this way the betweenness and metric properties of a line are simple, trivial consequences of the known properties of the real number system.

In addition to the expected material, various topics are introduced as optional, hopefully interesting, material for the reader. These include spherical geometry, the parallel postulate and absolute geometry, Desargues' Theorem, the Hilbert axioms, finite geometries, the integer-sided right triangles, cardinal number, and (inexplicably) the Jordan Curve Theorem.

* A more comprehensive review of this book will appear later.

This reviewer is left with a few quibbles. Suppose we accept the theory subscribed to by the author, namely that the early and systematic use of the real number system is the most appropriate approach for the readers of this text. Is it, nevertheless, fair to state that this "uses the all-important notion of a function in a way and to an extent that makes this approach seem relatively near to the main stream of mathematics"? Is this, then, the purpose of the ruler postulate? Also, should not something be done to relieve students of the thought that axioms are introduced by the merest whim of the mathematician? To say that ". . . a theorem that is excessively difficult or subtle can often be replaced by an extra postulate . . ." cannot give the naive reader too healthy an impression of what a postulate is all about. In this text, for example, the area axioms might have been given as a theorem which was too difficult to prove.

But quibbles aside, the text does succeed in its aims. It presents a coherent and intelligent geometry course to education majors with enough additional material to allow different teachers teach different courses.

MELVIN HAUSNER, New York University

Elements of Functional Analysis. By I. J. Maddox. Cambridge University Press, New York, 1970. 218 pp. $6.95. (Telegraphic Review, May 1970.)

Anyone teaching an undergraduate functional analysis course should consider this book as a possible text, along with others such as H. L. Royden, *Real Analysis*, G. F. Simmons, *Introduction to Topology and Modern Analysis*, Goffman-Pedrick, *First Course in Functional Analysis*, Lusternik-Sobolev, *Elements of Functional Analysis*, and also those discussed in the survey of potential undergraduate real analysis textbooks given in this Monthly, 75 (1968), 1033–1035. The present book aims to be more introductory in nature than those mentioned above. The seven chapters deal respectively with the elements of set theory (e.g., Zorn's Lemma), metric spaces (culminating in the uniform boundedness principle), linear spaces (including distributions defined on test functions of one variable), Banach spaces (including the Banach isomorphism theorems for operators), Banach Algebras (consisting essentially of the development of one version of the Gelfand representation theorem), Hilbert space (Riesz representation theorem), and matrices operating on sequence spaces. Clearly the author has exercised the well-known prerogative of writing the last chapter somewhat more in accordance with his own interests as a mathematician. On the other hand the emphasis on sequence spaces is consistent with the attempt to give examples without developing any measure or integration theory (although $L_p[0, 1]$ is introduced as an example of the equivalence of completeness and convergence of absolutely convergent series).

In an attempt to render some meaningful perspective, it seems to the reviewer that the following comparisons with the books mentioned above should perhaps be offered. This book probably does not read as well as that of Simmons; on the other hand in Simmons one must cover more material and more topological and algebraic preliminaries. Royden is more concise, but not as suitable

for a truly undergraduate level course. Like Simmons, Maddox avoids integration theory. Except for the lack of integration theory, the scope of his book is somewhat similar to that of Goffman-Pedrick, although of a much more limited nature. Also it contains neither self-adjoint operator theory nor nonlinear functional analysis which are treated, for example, in Lusternik-Sobolev.

In the reviewer's opinion the author has, with the exception of the last chapter, successfully adhered to his program of writing a book suitable for a "really introductory, though nontrivial, course on functional analysis for undergraduates." More than 300 exercises are supplied, many of which are truly exercises rather than additional theorems stated in problem format.

KARL GUSTAFSON, University of Colorado

FILMS

Newton's Method. By Herbert Wilf. Calculus Film Project of the MAA under the direction of H. M. MacNeille. Available (rent or buy) from Modern Learning Aids in the U. S. and Canada. 10 min., 16 mm., color.

In this film quadratic equations are fed into a multi-colored box which digests them, beeps several times, and then produces the roots. It is suggested that everyone should have such a box. Upon closer examination it is found that the box utilizes the quadratic formula. Then other types of equations are exhibited which the box is unable to digest. As a result a new multi-colored box is produced which takes these equations, together with a guess at a root, and produces a new number which is a better guess. Then the box accepts the better guess and produces an even better guess. The process continues until the result is as accurate as desired. This box is shown to utilize Newton's method which is then explained. Then the box finds a good approximation to the root of x^2-2, and the film concludes with some remarks about the iterative process while a new multi-colored box produces better and better approximations to "The End."

This is a good film to use when introducing Newton's method. It provides enough to enable the student to work with the idea but it doesn't show him everything. (The student may discover some of these things, including what can go wrong, if he has access to a sophisticated calculator or computer to help him with the calculations.) The film requires a knowledge of the derivatives of standard functions and an understanding that the slope of the tangent is given by the derivative.

While some of the calculus students who saw the film felt it was a little juvenile, most felt it was a good way to present Newton's method. All agreed that films of this nature were far better than most of the other mathematics films they had seen which usually lacked imagination and in which the sound was usually garbled.

DON KOEHLER, Miami University

A Function is a Mapping, Continuity of Mapping. Two films produced by the Committee of Educational Media of the MAA, supported by a grant from the National Science Foundation. Mathematician: Albert G. Fadell. Project Director: H. M. MacNeille. Distributed by Modern Learning Aids.

The first film can be used with profit in a pre-calculus course although both films are suitable for an elementary calculus course.

The length of each of these 16 mm films is 10 minutes, and both are done in color with an intriguing musical background. Concepts are presented by means of animation and narration. The animation in both films describes a function by moving representative points in the domain space to their images in the range space. This procedure gives a strong intuitive interpretation of functions which is especially helpful for introducing the concept of continuity.

The first film defines a function, in the narration, as a set of ordered pairs of elements no two of which have the same first element. The set of first elements is called the domain of the function while the set of second elements is called the range.

At the beginning of the film, no restriction is placed on the domain or range of a function, but the examples given are of real valued functions of a real variable. Special cases of the functions $K_c(x)=c$, $I(x)=x$, $T_c(x)=x+c$, $M_c(x)=cx$ are delineated by animation. The absolute value function is defined as the identity function on the right of zero and the negative function M_{-1} on the left of zero. The final example given is the inversion function sending x to $1/x$.

In the second film, the concepts of continuity and discontinuity are very effectively motivated by illustrating a function having the property that points near a are sent to images remote from the image of a.

A second example shows some points near a being sent near the image of a while other points near a have images remote from the image of a. Finally the idea of functions sending all points near a to images near the image of a is illustrated and it is explained that these special functions are said to be continuous at the point a.

The problem of what one means by "near" is settled by introducing the concept of open neighborhood. Both the open neighborhood and the ϵ, δ definitions of continuity are stated.

One theorem is proved, that being that the composite of two continuous functions is continuous. The animation for this theorem is especially effective.

Since the concepts presented here are often not easily grasped, these two films should be a welcome addition to the instructor's use of the text and classroom lecture. Although students are generally appreciative of these films on first showing, it has been observed that a second showing, after working with the concepts for several weeks, elicits greater enthusiasm and broader understanding.

S. R. Hilding, Gustavus Adolphus College

TELEGRAPHIC REVIEWS

Telegraphic reviews are intended to give prompt notice of books, with sufficient information to assist in deciding whether to order an examination copy or suggest library purchase. Possible uses are indicated as follows:

B = college bookstore stock   L = library purchase
P = professional reading   S = supplementary reading
T = textbook   E = teacher education
13 to 18 = freshman to second year graduate level usage
1 to 4 = approximate time in semesters to cover text
* = positive emphasis   ? = negative emphasis

Books on high school material (pre-calculus) are denoted REMEDIAL, and normally receive telegraphic reviews only if they are written for college students. Publishers are denoted by the standard abbreviations used in *Books in Print*, which gives complete addresses.

ALGEBRA, *T(14-15: 2), S, L, *Classical Modern Algebra.* Seth Warner. P-H, 1971, 520 pp, $14.95. Contents are drawn from the first eight chapters of the author's *Modern Algebra.* Material on finite groups, including the Sylow theorems, has been added. Not enough material for a thorough course in linear algebra. A large number of exercises. L.C.L.

ALGEBRA, INTRODUCTORY, T+++(13-14: 1), E, S, *Foundations of Algebra and Number Theory.* John M. Peterson. Markham, 1971, 142 pp, $8.50. A leisurely, but quite careful, approach to a modern algebra course for school-teachers. Sets, relations, and Number Theory are introduced and then the Reals are obtained through Dedekind Cuts. Linear systems are done at length. Fundamental definitions and axioms are given for rings and groups. W.C.R.

ALGEBRA, LOGIC, P, L, *Sixteen Papers on Logic and Algebra. American Mathematical Society Translations, Series 2, Volume 94.* AMS. AMS, 1970, 276 pp, $14. Topics from the theory of algorithms, symbolic logic, model theory, lattice theory, abelian groups, and semigroups. L.C.L.

ALGEBRAIC GEOMETRY, P, L, *Selected Topics in Algebraic Geometry.* Virgil Snyder, et al. Chelsea, 1970, 484 pp, $8.75 (two volumes in one). A reprint, in one volume, with corrections of errata, and references in the bibliography to new editions of books and to collected works, of the Report of the Committee on Rational Transformations, Division of Physical Sciences, National Research Council, originally published in 1928 and 1934 as Numbers 63 and 96 of the Bulletin of the NRC. No index, but a very large bibliography, broken down by subject matter. J.D.-B.

ANALYSIS, T(17-18), S, P, L, *Complementary Variational Principles.* A.M. Arthurs. Oxford U Pr, 1970, 95 pp, $6.50. An introduction to the theory and applications of complementary variational principles, assuming an elementary knowledge of calculus of variations and linear operator theory. After sufficient introductory material, the author considers applications to the theory of linear and non lineary boundary value problems. Contains an extensive bibliography. T.A.V.

APPLICATIONS FOR CHEMISTS, *T(16-17: 1), S, B, L. *Mathematics for Chemists*. Charles L. Perrin. Wiley, 1970, 453 pp, $11.95. Beginning with a review of calculus, there follows useful methods from analysis, probability, linear algebra (finite and infinite), and group theory. The examples are from physical chemistry. Meant as a text, there are many exercises. There is no attempt at rigor. W.C.R.

APPLIED ANALYSIS, T(13-14), S. *A Course of Mathematics for Engineers and Scientists (vol. 1, 2nd ed)*. Brian H. Chirgwin and Charles Plumpton. Pergamon Pr, 1970, 548 pp, $6.75. Emphasis on examples and exercises to illustrate methods, techniques and applications of differentiation and integration of one variable, analytic geometry of two dimensions and elementary complex number theory. Substantial increase of numerical work from the first edition, plus a new chapter on elementary probability and statistics. L.C.L.

BIOGRAPHY. *Quelques Aspects de La Pensée D'Un Mathématicien*. Paul Levy. Albert Blanchard, 1970, 222 pp, $6.18 (P). An autobiography with mathematical content and a second part on his philosophy in general. J.A.S.

CALCULUS, **T(13: 2). *Introduction to the Calculus*. I.N. Herstein and Reuben Sandler. Har-Row, 1971, 309 pp, $. An intuitive approach to the calculus which is honest and correct without being sloppy, by limiting the treatment to well-behaved functions, problems which are not tricky, clever, or cute, and staying with the one-variable case. The book has been kept small (285 pages). A look at his skillful introduction to complex numbers in order to develop an interrelationship between hyperbolic functions and trigonometric ones, gives one a feeling for the beauty of this book. A desk copy is a must, even if a teacher prefers a text with more rigor; each presentation is a gem which could give some insight into techniques of lecturing. L.L.K.

CALCULUS, S(13-14: 1). *Calculus Supplement. An Outline With Solved Problems*. Robert Kurtz. Benjamin, 1970, 274 pp, $3.95 (P). A concise outline of calculus, covering the derivative, antiderivative, integral and infinite series, with exercises and problems (149 pages) and solutions (120 pages). Designed for self-teaching or review and would take a great deal of effort on the part of a teacher if used as a text. L.L.K.

CATEGORY THEORY, *P, L. *Kan Extensions in Enriched Category Theory. Lecture Notes in Mathematics, No. 145*. Eduardo J. Dubuc. Springer-Verlag, 1970, 172 pp, $4.40 (P). A timely and efficient exposition of the important recent work on enriched categories. Includes a helpful section on terminology. J.A.S.

COHOMOLOGY (RINGS), COHOMOLOGICAL DIMENSION, P, L. *Cech Cohomological Dimensions for Commutative Rings. Lecture Notes in Mathematics, No. 147*. David E. Dobbs. Springer-Verlag, 1970, 176 pp, $4.60 (P). The author's intent is to "suggest some connections between Cech cohomology and the dimension theories arising from Grothendieck cohomology that have been extensively studied by Artin and Grothendieck." Some of the results may also relate to Amitsur cohomology. Contains a serious attempt to present some of the folklore of the subject. J.A.S.

COMBINATORICS, P, *L. *Graph Theory and Its Applications.* Bernard Harris. Acad Pr, 1970, 262 pp, $5. The proceedings of an advanced seminar conducted by the Mathematics Research Center, U.S. Army, at the University of Wisconsin, October 1969. A timely work containing eleven papers on combinatorics and graph theory from a wide variety of views. The diversity of the papers shows the active research interest in the field in many disciplines. T.A.V.

COMBINATORICS, P, L. *On the Foundations of Combinatorial Theory: Combinatorial Geometries (Preliminary edition).* Henry H. Crapo and Gian-Carlo Rota. M.I.T. Pr, 1970, 328 pp, $10 (P). An exposition, involving some original work, of the current state of combinatorial geometry. Extensive bibliography. J.A.S.

CONTINUOUS TRANSFORMATION GROUPS, P, L. *Theorie der Transformationsgruppen.* Sophus Lie. Chelsea, 1970, $49 (3 vols). *Vol. I*, 645 pp; *Vol. II*, 568 pp; *Vol. III*, 830 pp. A reprint, with errata corrected and a subject index for all three volumes added at the end of each of the first two, of Lie's classic treatise on continuous transformation groups, first published, in three volumes, in 1888, 1890, and 1893. J.D.-B.

DIFFERENTIAL EQUATIONS, T(14: 1). *Ordinary Differential Equations. A Programmed Course for Students of Science and Technology.* Ed: A.C. Bajpai, I.M. Calus and J. Hyslop. Wiley, 1970, 249 pp, $7.50 (P). This is a subject which seems particularly suitable for the programmed method of presentation and this text seems to live up to the authors claim: a series of programs developed for scientists interested in solution for the types of equation which occur most frequently in physical problems, with more theoretical aspects omitted. Covers first order differential equations, differential equations with constant coefficients (solutions by D-operator methods and by Laplace Transform methods), simultaneous differential equations, and linear differential equations; all by using a three step presentation of recognition, techniques of solution, and applications. L.L.K.

ELECTRICITY, REPRINT, P, L. *Electrical Papers, Vols. I and II.* Oliver Heaviside. Chelsea, 1970, 560 pp and 587 pp, $29.50 set. An exact reprint of the two-volume *Electrical Papers* of Heaviside published in 1892 by Macmillan, except for correction of errata. (Another reprint of the 1892 edition, containing a biographical sketch of Heaviside by Sir Oliver Leach, was published in 1925 by The Copley Publishers on the event of Heaviside's death.) Papers on telegraphy, propagation of variations of current along wires, electrical theory in general, and electromagnetic induction and its propagation are included in this collection. D.F.A.

FUNCTIONS OF A COMPLEX VARIABLE, *T(15-17: 1, 2), L. *Complex Variables Applied in Science and Engineering.* Harold Wayland. Van Nostrand, 1970, 350 pp, $9.50 (P). An introduction to the theory of functions of a complex variable for students of science and engineering, with application of this theory to the evaluation of definite integrals by using the theory of residues and to the solution of two-dimensional potential theory problems. Important parts of the book are the further application of the theory to the solution of the common second-order linear differential equations of mathematical physics, and a study of the special functions

associated with these equations. D.F.A.

GENERAL, *T(13: 1, 2; NON-MATH). *Finite Mathematics.* Francis H. Hildebrand and Cheryl G. Johnson. Prindle, 1970, 407 pp, $9.95. Topics: logic, vectors and matrices, probability, statistics, linear programming, theory of games. Appendices deal with $\sum$ notation, powers and roots, and binomial, standard normal and chi-square distributions. J.N.C.

GENERAL APPLIED MATHEMATICS, T(14-16: 1, 2), L. *An Introduction to Industrial Mathematics.* M. Rothman. Van Nostrand, 1970, 370 pp, $18.50. A cursory look at many topics in mathematics which are encountered by engineers. Although there are many references for further study, there are no exercises. Perhaps useful as an engineering mathematics text or for self-study by a scientist in industry. Some of the topics: Fourier series, differentiation, integration, the Laplace transform, statistics, numerical analysis, operational research, vector analysis, mechanics. D.F.A.

GENERAL TOPOLOGY, P, L. *Lectures on Topological Dynamics.* Robert Ellis. Benjamin, 1969, 211 pp, $17.50. A volume in Benjamin's *Mathematics Lecture Note Series*, these self-contained notes are intended to give a unified presentation of recent results in topological dynamics and to describe research techniques which have proved useful. The main theme of the book is the classification of minimal sets, and an algebraic approach is taken; useful references, notes, and mention of unsolved problems are provided. The book is printed directly from typescript. D.F.A.

GEOMETRY, *S, *P, *L. *Miniquaternion Geometry: An Introduction to the Study of Projective Planes.* T.G. Room and P.B. Kirkpatrick. Cambridge U Pr, 1971, 176 pp, $12.50. Number 60 in Cambridge Tracts in Mathematics and Mathematical Physics, this volume introduces four finite geometrical systems and the subject of projective planes. Three of these geometries are constructed from the miniquaternion system by using different methods of coordinatization and the fourth is obtained from the field of order 9. Incorporates numerous valuable exercises varying in difficulty. Assumes a first course in abstract algebra. J.N.C.

GEOMETRY, L. *Bibliography of Non-Euclidean Geometry.* D.M.Y. Sommerville. Chelsea, 1970, 403 pp, $12. A supplemented reprint of a work originally published in 1911 under the title *Bibliography of Non-Euclidean Geometry, Including the Theory of Parallels, The Foundations of Geometry, and Space of n Dimensions.* J.N.C.

GEOMETRY, E(1, ELEMENTARY). *Modern Elementary Geometry.* James M. Moser. P-H, 1971, 333 pp, $8.95. Prepared along the guidelines of the CUPM in its Level I recommendations for the training of teachers. Begins with informal geometry and later examines more formal aspects of geometry. J.N.C.

GEOMETRY, S, L. *Ruler and the Round or Angle Trisection and Circle Division, Volume 15.* Nicholas D. Kazarinoff. Prindle, 1970, 138 pp, $3.50 (P). Another in Prindle's Complementary Series in Mathematics. J.N.C.

GEOMETRY, E(1, JH & SH), S, L. *Plane Geometry: An Approach Through Isometries*. Dick W. Hall and Steven Szabo. P-H, 1971, 209 pp, \$8.95. Isometries are introduced and used to prove some of the more elementary theorems of Euclidean geometry. The parallel postulate is not presented until Chapter 9. Appendices list the axioms, definitions and theorems of the text. J.N.C.

GEOMETRY, T(15-16), L. *A Guide to Undergraduate Projective Geometry* A.F. Horadam. Pergamon Pr, 1970, 349 pp, \$9.50. Contains a 43 page introduction which describes interesting properties of the projective plane and their historical development. The remainder is a systematic treatment of projective geometry with emphasis on finite geometry, and the connection between geometries and groups of transformations. J.N.C.

HANDBOOK, B. *Engineering Mathematics Handbook*. Jan J. Tuma. McGraw, 1970, 334 pp, \$9.95. Subtitle - *Definitions, Theorems, Formulas, and Tables* - "a summary of the major tools of engineering mathematics...for engineers, scientists, and architects... A pictorial dictionary." L.A.S.

HISTORY. *Rara Arithmetica*. David E. Smith. Chelsea, 1970, 725 pp, \$15. A catalog for bibliophiles of arithmetic books prior to 1601, in the Plimpton collection at Columbia first published 1908. Also reprinted in this edition is the volume DeMorgan's *Arithmetical Books* (1847). L.A.S.

HISTORY, S, P, *L. *Jan Lukasiewicz, Selected Works*. Ed: L. Borkowski. North-Holland, 1970, 405 pp, \$20. These papers have been selected to bring out Lukasiewicz's interest in the problem of determinism. They include the first formulation of his systems of many-valued logic and his pioneer works in the field of the methodology of propositional calculus and the history of logic, as well as his views on a number of philosophical and logical issues. L.C.L.

HISTORY OF LOGIC, P, L. *A History of Formal Logic*. I.M. Bochenski. Transl and Ed: Ivo Thomas. Chelsea, 1970, 567 pp, \$11.50. A collection of excerpts, with commentary by the author, which give contemporary views of problems in formal logic at major periods of development, Greek, Scholastic, Indian and modern (mathematical) logic (to 1930). Massive historical and philosophical bibliography. L.A.S.

HOMOLOGY THEORIES, P, L. *Unoriented Bordism and Actions of Finite Groups. Memoirs of the American Mathematical Society, Number 103*. R.E. Stong. AMS, 1970, 80 pp, \$1.80 (P). An analysis of the equivariant homology theories arising from equivariant unoriented bordism. J.A.S.

LINEAR PROGRAMMING, L. *An Illustrated Guide to Linear Programming*. Saul I. Gass. McGraw, 1970, 173 pp, \$9.95. Written for a general audience - "without the usual involved mathematical complexities". An interesting sequence of examples, anecdotes and illustrations introduces problem formulation. The only method detailed is the simplex method in two dimensions. Applications include network, traveling salesman, contract awarding and game theory problems. An appendix lists typical terms, techniques and applications. R.W.N.

LOGIC, P, S. *Frege and Gödel: Two Fundamental Texts in Mathematical Logic*. Ed: Jean Van Heijenoort. Harvard U Pr, 1970, 116 pp, $6. These English translations of Frege's *Begriffschrift* (1879) and Gödel's incompleteness paper (1931) were originally published in *From Frege to Gödel: A Source Book in Mathematical Logic, 1879-1931*, same editor and publisher, 1967. Those who cannot afford the expanded volume will welcome this opportunity to obtain these two important translations. K.W.

MARKOV CHAINS, T(16-17), S. *Lecture Notes in Operations Research and Mathematical Systems: Markovketten-35*. F. Ferschl. Springer-Verlag, 1970, 168 pp, $3.90 (P). An expansion of the lecture notes for a course at the University of Bonn taken chiefly by students of economics, this book is an introduction to the theory of Markov chains with denumerable state spaces. The mathematical prerequisites are not extensive, but notions as sophisticated as the convergence of doubly-infinite series and Cesàro summability are introduced and used. J.D.-B.

MATHEMATICAL PHYSICS, P, L. *Vector Bundles in Mathematical Physics, Volume 1*. Robert Hermann. Benjamin, 1970, 441 pp, $17.50. An apparently successful attempt to open the world of modern differential geometry to physicists and to develop some mathematically interesting physical interpretations in the general area of quantum mechanics. Develops and uses the theory of jet bundles. J.A.S.

MATHEMATICAL PHYSICS, T(14-16: 2), **S, B, L. *Lectures in Mathematical Physics, Volume 1*. Robert Hermann. Benjamin, 1970, 475 pp, $17.50. Vector spaces, classical mechanics, differential equations, Hilbert space, and even generalized functions nicely presented in typescript. A bit expensive considering the mathematics lecture note series. Professor Hermann really appears to achieve the exciting interdisciplinary effect that is the aim of this novel book. Beautiful coverage of folklore and intuitative background. No exercises. J.A.S.

MATHEMATICAL PHYSICS, P, L. *Fourier Analysis on Groups and Partial Wave Analysis*. Robert Hermann. Benjamin, 1969, 302 pp, $17.50. One of the books in Benjamin's *Mathematics Lecture Note Series*, this volume examines mathematical problems arising in elementary particle physics, and its intended audience includes mathematicians and physicists. Topics include meromorphic decompositions and analytic continuation, the Fourier transform on Lie groups, Cauchy integrals on Lie groups, partial wave analysis as a problem in group representation theory, and generalized functions on manifolds. The book is printed directly from typescript, and contains many bibliographic references. D.F.A.

MATHEMATICAL PHYSICS, S, *P, L. *Lie Algebras and Quantum Mechanics*. Robert Hermann. Benjamin, 1970, 320 pp, $17.50. A series of expository papers with bibliographies on the current state of the active interfaces of physics and mathematics. Assumes a comfortable acquaintance with differential geometry, Lie groups, and vector bundles as well as quantum mechanics, current algebras, and cohomology theory. A delightful introduction in which some philosophy on the relations (hoped for) between physics and mathematics sets the tone of this volume aimed more at a mathematician than a physicist. J.A.S.

NUMBER THEORY, *T(15-16: 1), S, *L. *Elementary Number Theory: An Algebraic Approach.* Ethan D. Bolker. Benjamin, 1970, 180 pp, $10.50. The exposition is tied to the study of three classical problems: The structure of the group of units of Z_n, integers representable in the form x^2-my^2 and the Fermat equation $x^n + y^n = z^n$ for n = 2, 3 and 4. Excellent set of challenging problems. A valuable supplement for beginning courses in modern algebra and algebraic number theory. L.C.L.

OPERATIONAL CALCULUS, S, P, L. *Heaviside Operational Calculus: An Elementary Foundation.* Douglas H. Moore. Volume 30, *Modern Analytic and Computational Methods in Science and Mathematics.* Am Elsevier, 1971, 152 pp, $16. A reference book intended for teachers of undergraduate mathematics and engineering courses which provides a basis for Heaviside operational calculus. Material provided is accessible to undergraduates and could be useful as supplementary reading for them, since it includes many examples and applications of the theory to engineering problems. Book contains an historical foreward on Heaviside by E.T. Whittaker. D.F.A.

OPTIMIZATION, S, P, L. *Approximate Methods in Optimization Problems.* Vladimir F. Demyanov and Aleksandr M. Rubinov. Am Elsevier, 1970, 256 pp, $18.50. A translation and revision of the 1968 Russian edition. A summary of functional analysis enables the expression of necessary conditions in terms of cones. Conditional-gradient and gradient projection methods are analyzed. Considers some optimal control and finite dimensional problems. Nearly minimal index.R.W.N.

ORDINARY DIFFERENTIAL EQUATIONS, S, P, L. *Solution of Non-Linear Systems.* Stanislav Vojtášek and Karel Janáč. Daniel Davey, 1970 (c. 1969), 246 pp, $13.50. An introduction to some of the qualitative theory of non-linear systems, with discussion of what various kinds of stability are, and an explanation of mathematical methods of solution of non-linear problems and of means of solution of them using analogue computers. Possibilities for solution by digital computers are briefly discussed, and six examples of practical solutions are presented. The book is designed to be useful to someone interested in theory or in practical methods of solution, or both. Translated from Czech. D.F.A.

PROBLEMS, S, P, B. *Problèmes De Probabilité.* Gérard Letac. Presses Universitaires De France. 1970, 118 pp, 12F. 74 non trivial problems and solutions in probability theory useful in a strong first or second course. Useful international bibliography to French problems. J.A.S.

REMEDIAL, T(13: 1). *Introduction to the Elementary Functions.* Ronald A. Knight and William E. Hoff. Dickenson, 1969, 263 pp, $7.95. Designed as a one semester pre-calculus course in the study of elementary functions: exponentials; logarithms; power, polynomial and rational functions; trigonometry; and a chapter on inequalities, absolute values and mathematical induction. Included in the text is a list of symbols, some tables, and a glossary. This course could remove some of the pressures from the first year calculus if taught in the high schools. L.L.K.

SEMI-MARKOV PROCESSES, P, L. (RESEARCH). *Lecture Notes in Operations Research and Mathematical Systems: Semi-Markoff Prozesse mit endlich*

vielen Zuständen-34. H. Störmer. Springer-Verlag, 1970, 128 pp, $3.30 (P). A sketch of the theory of renewal processes, followed by a derivation of some of the results on semi-markov processes with finite state spaces which are most useful in applications. J.D.-B.

SEVERAL COMPLEX VARIABLES, S, *P, *L, *Theorie der Funktionen mehrerer komplexer Veränderlichen.* H. Behnke and P. Thullen. Springer-Verlag, 1970, 225 pp, $13.20. This second edition of the earlier (1934) classic on the theory of functions of several complex variables has been updated by the inclusion of a selection of more recent developments in the field. In addition, a very extensive bibliography has been included. T.A.V.

TOPOLOGY, *S, T, *P, *L, *Counterexamples in Topology.* Lynn A. Steen and J. Arthur Seebach, Jr. HR & W, 1970, 210 pp, $8. Thirty-eight pages of basic definitions (including separation axioms, compactness, connectedness, metric spaces), one hundred and forty-three counterexamples, six reference charts displaying various properties of the examples, one hundred and forty-seven problems, twenty-three pages of notes, and a basic bibliography. The book appears to live up to its excellent first sentence: "The creative process of mathematics, both historically and individually, may be described as a counterpoint between theorems and examples." K.O.M.

STATISTICS, TABLES, P, L, *Selected Tables in Mathematical Statistics, Volume I.* H.L. Harter and D.B. Owen. Markham, 1970, 405 pp, $5.80. First of a possible series of statistical tables sponsored by The Institute of Mathematical Statistics. Contains four new sets of tables plus an extension of the tables of critical values and probability levels for the Wilcoxon Rank Sum Test and the Wilcoxon Signed Rank Test. Photo-offset printing. R.S.K.

Reviewers Whose Initials Appear Above

David F. Appleyard, Carleton; Judith N. Cederberg, St. Olaf; John Dyer-Bennet, Carleton; Lorraine L. Keller, St. Olaf; Loren C. Larson, St. Olaf; Kenneth O. May, University of Toronto; R.W. Nau, Carleton; William C. Ramaley, Carleton; J. Arthur Seebach, Jr., St. Olaf; Linda A. Seebach, St. Olaf; T.A. Vessey, St. Olaf; Kenneth Wegner, Carleton.

NOTABLE

Among the above reviews there are four classified as MATHEMATICAL PHYSICS. Two of these books are the first volumes of proposed two volume sets. Without having seen these remaining two volumes it seems safe to say that these six books represent a significant and most likely successful attempt at a very extensive expository effort in the area of modern mathematics and its (potential) relations to physics. We feel it is of great value to have available this series of books relating some of the most significant developments of 20th century mathematics to theoretical physics at a level that should be of interest to professionals in both disciplines. All of the volumes are in the Benjamin Mathematics Lecture Notes Series and apparently are also available in paperback at $7.95 each.

NEWS AND NOTICES

EDITED BY RAOUL HAILPERN, SUNY at Buffalo

Readers are invited to contribute to the general interest of this Department by sending news items to Mathematical Association of America, 1225 Connecticut Avenue, N.W., Washington, D. C. 20036. Items must be submitted at least two months before publication can take place.

PERSONAL ITEMS

Dean D. E. Dupree, Northeast Louisiana University, represented the Association at the inauguration of J. H. Allen as 32nd President of Centenary College of Louisiana on October 21, 1970.

Professor Francis Regan, St. Louis University, was awarded the C. C. MacDuffee Award for Distinguished Service in Mathematics.

East Texas State University: Dr. Charles Wall, University of Tennessee, has been appointed Assistant Professor; Mr. Joseph Spacek, University of South Alabama, has been appointed Instructor of Computer Science.

University of Georgia: Dr. Myron Rosskopf, Teachers College, Columbia University, has been appointed Visiting Professor in the Department of Mathematics Education; Drs. Edith Robinson and Leonard Pikaart are postdoctoral fellows at New York University for the academic year 1970–71.

Lake Forest College: Dr. Dianne Salam, Texas Christian University, has been appointed Instructor; Associate Professor R. J. Troyer has been promoted to Professor.

State University College at Fredonia: Drs. Gerald Giaccai, University of Illinois, and Albert Polimeni, University of Syracuse, have been appointed Assistant Professors; Dr. Eugene Rozycki, SUNY at Binghamton, has been appointed Associate Professor.

Winona State College: Associate Professor H. A. Heckart, South Dakota School of Mines and Technology, has been appointed Associate Professor and Head of the Department of Mathematical Sciences; Assistant Professor Marceline Gratiaa is on sabbatical leave.

Associate Professor F. B. Allen, Elmhurst College, has been appointed Chairman of the Department of Mathematics.

Assistant Professor W. R. Boland, Drexel University, has been appointed Assistant Professor at Clemson University.

Mr. M. J. Brown, Northern Kentucky State College, has been promoted to Assistant Professor.

Mr. D. K. Buck, Wytheville Community College, has been appointed Associate Professor at Tennessee Technological University.

Mr. C. P. Campbell, West Virginia State College, has been promoted to Assistant Professor.

Dr. J. W. Carlson, University of Missouri at Rolla, has been appointed Assistant Professor at Kansas State Teachers College.

Dr. S. D. Chatterji, formerly of the University of Copenhagen, has accepted the position of "professeur ordinaire" at the Ecole Polytechnique Fédérale of Lausanne.

Assistant Professor Hugh Coomes, SUNY at Albany, has been appointed Associate Professor at Paterson State College.

Dr. R. A. Cooper, Texas Technological University, has been appointed Assistant Professor at Trinity University.

Professor R. E. Doutt, South Dakota School of Mines and Technology, retired in July 1970 with the title of Professor Emeritus.

Assistant Professor W. L. Drezdzon, Chairman of the Mathematics Department at Kennedy King College, has been promoted to Associate Professor.

Dr. R. H. Dumonceaux, St. John's University, has been appointed Chairman of the Mathematics Department.

Assistant Professor A. R. Elcrat, Wichita State University, has been promoted to Associate Professor.

Assistant Professor John Firkins, Gonzaga University, has been promoted to Associate Professor; he was selected in 1969 as the "Outstanding Teacher of the Year" by students and faculty at Gonzaga University. He has also been named to serve a three year term on the Teacher Education Liaison Committee for the Washington State Department of Education.

Associate Professor R. F. Gundy, Rutgers University, has been promoted to Professor.

Dr. Chung-wu Ho, MIT, has been appointed Assistant Professor at Southern Illinois University.

Mr. Edwin Hoefer, SUNY at Buffalo, has been appointed Assistant Professor at Rosary Hill College.

Professor R. W. Hunt, U. S. Naval Postgraduate School and Southern Illinois University, has been appointed Professor and Head of the Mathematics Department at California State College, Bakersfield.

Mr. M. A. Johnson, Ashland College, has been promoted to Assistant Professor.

Associate Professor L. S. Kennison, Brooklyn College, has been appointed Professor at Southeastern Massachusetts University.

Dr. James Lepowski, MIT, has been appointed Research Instructor at Brandeis University.

Assistant Professor E. J. Montella, Merrimack College, has been promoted to Associate Professor.

Professor Nand Kishore, University of Toledo, died on November 28, 1969 at the age of 56. He was a member of the Association for seven years.

Dr. G. B. Parrish, U. S. Army, Research Office, Durham, died on December 25, 1969 at the age of 41. He was a member of the Association for thirteen years.

Assistant Professor G. T. Roberts, Caldwell College, died on December 19, 1969 at the age of 33. He was a member of the Association for four years.

Reverend J. W. Sehestedt, Stigler, Oklahoma, died on March 6, 1970 at the age of 76. He was a member of the Association for eighteen years.

Professor H. S. Zuckerman, University of Washington, died on June 16, 1970 at the age of 58. He was a member of the Association for twenty-six years.

PRELIMINARY ANNOUNCEMENT: 13TH BIENNIAL SEMINAR OF THE CANADIAN MATHEMATICAL CONGRESS

The 13th Biennial Seminar of the Canadian Mathematical Congress will be held at Dalhousie University, Halifax, Nova Scotia, August 16–September 3, 1971, inclusive. The theme will be, "Differential Geometry, Differential Topology and Applications."

The program of lectures of the Seminar is now being completed, and a brochure describing the lecture program and facilities at Dalhousie University will be sent to Universities and Research Centers. Three or four lecturers will each give six lectures on their recent research in topics within the field of the seminar; three or four other lecturers will each give about twelve lectures on topics related more or less closely with the themes of the more condensed lecture series. In addition, about eight to ten other specialists in the field of the seminar will be invited to participate in the seminar for a few days and to give one or two lectures. Past experience suggests that the participants will create a number of *ad hoc* seminar series on topics which interest them.

It is hoped that 80 to 100 participants may attend the seminar, and that full or partial support may be available to some participants. A tear-off application form will form part of the descriptive brochure mentioned above. Mathematicians interested in the theme

of the seminar are invited to write either to the secretariat at the address below for a copy of the brochure or to the Chairman of the Program Committee.

Program Committee: J. R. Vanstone (Chairman), University of Toronto; A. E. Fekete, Memorial University of Newfoundland; W. H. Greub, University of Toronto; H. Rund, University of Waterloo; U. Suter, Forschung Institut für Mathematik, Zürich, Switzerland; N. Van Que, Université de Montréal.

Secretariat: Canadian Mathematical Congress, 985 Sherbrooke Street West, Montreal 110, Quebec, Canada.

MATHEMATICAL ASSOCIATION OF AMERICA

Official Reports and Communications

JUNE MEETING OF THE PACIFIC NORTHWEST SECTION

The Annual Meeting of the Pacific Northwest Section of the MAA was held at Pacific Lutheran University, Tacoma, Washington, June 18 and 19, 1970 in conjunction with the Six-Hundred-Seventy-Sixth Meeting of AMS and the annual meeting of the Northwest Section of SIAM. One-hundred-twenty persons were in attendance.

At the business meeting the following officers were elected: Chairman, Charles Curtis, University of Oregon; First Vice-Chairman, Maurice Kingston, University of Washington; Second Vice-Chairman, Lawrence Mitchell, Blue Mountain Community College; Secretary-Treasurer, Eugene Maier, University of Oregon.

The program of the MAA portion of the meeting was as follows:

1. *Why Contour Integration Leads to a Healthy Dose of Topology in Modern Differential Geometry*, by Richard Koch, University of Oregon.
2. *A Panel Discussion on the Purpose and Content of Upper Division Courses in Geometry*, Clark Benson, University of Oregon; J. V. Leahy, University of Oregon; T. G. Ostrom, Washington State University; J. R. Reay, Western Washington State College.
3. *The Lack of Mathematical Vitamins in the Development of Artificial Intelligence*, by Emilio Gugliardo, Oregon State University.
4. *Formal Logic: Its Curricular Relationship to Mathematical Studies*, by Walter Coale, Skagit Valley College.
5. *A Weighted Derivative*, by Richard Plagge, Highline Community College.
6. *Some Topics in Computer Science*, by Robert Arend, Spokane Community College.
7. *The In-Service Training Program for Oregon Community College Mathematics Teachers*, by Lawrence Mitchell, Blue Mountain Community College.
8. *A Demonstration of Dial Access Computer Systems*, by Guy Benton, Green River Community College and James Warjone, Service Bureau Corporation.
9. *Some Ideas on the Solution of Inequalities*, by Elmer Zimgalis, Highline Community College.
10. *Mathematics for Vocational-Technical Programs*, by Galen Nielson, Linn-Benton Community College.
11. *Some Ideas for Using the Computer in Freshman Mathematics*, by James Relf, Bellevue Community College.

E. A. MAIER, *Secretary*

CALENDAR OF FUTURE MEETINGS

Fifty-second Summer Meeting, Pennsylvania State University, University Park, August 30–September 1, 1971.

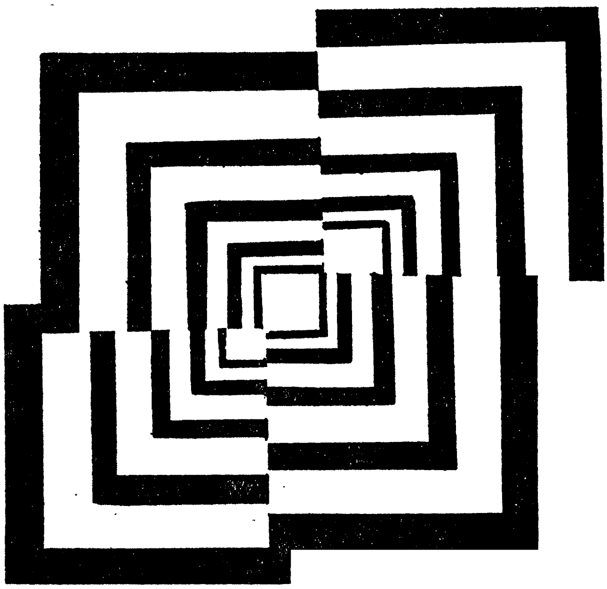
Fifty-fifth Annual Meeting, Las Vegas, Nevada, January 19–21, 1972.

The following is a list of the Sections of the Association with dates of future meetings so far as they have been reported to the Editorial Director.

ALLEGHENY MOUNTAIN, Geneva College, Beaver Falls, Pennsylvania, May 7–8, 1971.
FLORIDA
ILLINOIS, Eastern Illinois University, Charleston, May 14–15, 1971.
INDIANA, Purdue University, North Central Campus, Westville, May 8, 1971.
IOWA, Loras College, Dubuque, April 23, 1971.
KANSAS
KENTUCKY, Western Kentucky University, Bowling Green, April 2–3, 1971.
LOUISIANA-MISSISSIPPI
MARYLAND-DISTRICT OF COLUMBIA-VIRGINIA
METROPOLITAN NEW YORK, Nassau Community College, Long Island, April 3, 1971.
MICHIGAN, Western Michigan University, Kalamazoo, May 7–8, 1971.
MISSOURI, Missouri Southern College, Joplin, April 30–May 1, 1971.
NEBRASKA, Nebraska Wesleyan University, Lincoln, April 30–May 1, 1971.
NEW JERSEY
NORTH CENTRAL, University of Minnesota, Minneapolis, May 8, 1971.
NORTHEASTERN, Colby College, Waterville, Maine, June 19, 1971.
NORTHERN CALIFORNIA
OHIO, Ohio Wesleyan University, Delaware, April 30–May 1, 1971.
OKLAHOMA-ARKANSAS
PACIFIC NORTHWEST, Oregon State University, Corvallis, June 18–19, 1971.
PHILADELPHIA, Lafayette College, Easton, November 20, 1971.
ROCKY MOUNTAIN, Weber State College, Ogden, Utah, May 7–8, 1971.
SOUTHEASTERN
SOUTHERN CALIFORNIA
SOUTHWESTERN, Arizona State University, Tempe, April 2–3, 1971.
TEXAS, Midwestern University, Wichita Falls, April 16–17, 1971.
UPPER NEW YORK STATE, St. Lawrence University, Canton, May 8, 1971.
WISCONSIN, Ripon College, Ripon, April 30–May 1, 1971.

FUTURE MEETINGS OF OTHER ORGANIZATIONS

AMERICAN ASSOCIATION FOR THE ADVANCEMENT OF SCIENCE, Philadelphia, December 26–31, 1971.
AMERICAN MATHEMATICAL SOCIETY, Pennsylvania State University, University Park, August 31–September 3, 1971.
AMERICAN SOCIETY FOR ENGINEERING EDUCATION, U. S. Naval Academy, Annapolis, June 21–24, 1971.
ASSOCIATION FOR COMPUTING MACHINERY, Chicago, August 3–5, 1971.
ASSOCIATION FOR SYMBOLIC LOGIC
CENTRAL ASSOCIATION OF SCIENCE AND MATHEMATICS TEACHERS, Detroit, Michigan, November 25–27, 1971.
FIBONACCI ASSOCIATION, University of San Francisco, April 24, 1971.
INSTITUTE OF MATHEMATICAL STATISTICS, Fort Collins, Colorado, August 23–26, 1971.
MU ALPHA THETA, Pennsylvania State University, University Park, September 1, 1971.
NATIONAL COUNCIL OF TEACHERS OF MATHEMATICS, Anaheim, California, April 14–17, 1971.
OPERATIONS RESEARCH SOCIETY OF AMERICA, Sheraton Dallas, Dallas, May 5–7, 1971.
PI MU EPSILON, Pennsylvania State University, University Park, August 31–September 1, 1971.
SOCIETY FOR INDUSTRIAL AND APPLIED MATHEMATICS, Seattle, Washington, June 28–30, 1971.



GEORGE BANTA COMPANY, INC., MENASHA, WISCONSIN

THE AMERICAN MATHEMATICAL MONTHLY

(FOUNDED IN 1894 BY BENJAMIN F. FINKEL)

THE OFFICIAL JOURNAL OF

THE MATHEMATICAL ASSOCIATION OF AMERICA, INC.

VOLUME 78 NUMBER 4

CONTENTS

(Continued on inside cover)

APRIL 1971

NOTICE TO AUTHORS

Specialized research is usually unsuitable; see *Statement of Policy* (vol. 76, p. 2). Manuscript preparation: Please use the *Manual for Monthly Authors* (vol. 78, p. 1) and follow the format in current issues of the MONTHLY. Manuscripts should be typewritten, triple-spaced with wide margins; submit two copies and keep one for protection against loss. Each author receives 50 free reprints.

Backlog: Main Articles 9 months, Math. Notes 9 months, Research Problems 7 months, Classroom Notes 8 months, Math. Education 6 months.

EDITORIAL CORRESPONDENCE AND MAIN ARTICLES: to HARLEY FLANDERS, American Mathematical Monthly, Tel Aviv University, Ramat Aviv, Israel (see Notice, vol. 77, 1970, p. 555); NOTES, etc.: to the corresponding Associate Editor; ADVERTISING CORRESPONDENCE: to RAOUL HAILPERN, Mathematical Association of America, SUNY at Buffalo, Buffalo, N. Y. 14214; CHANGE OF ADDRESS and SUBSCRIPTIONS: to A. B. WILLCOX, Mathematical Association of America, 1225 Connecticut Ave., N.W., Washington, D. C. 20036.

Annual dues for members of the Association (including a subscription to the American Mathematical Monthly) are $12.50. For nonmembers the subscription price is $18.00.

PUBLISHED BY THE ASSOCIATION at Washington, D. C., and Menasha, Wisconsin during the months of January, February, March, April, May, June–July, August–September, October, November, December.

Second-class postage paid at Washington, D. C., and additional mailing offices. Acceptance for mailing at special rate of postage provided for in the Act of February 28, 1925, embodied in Paragraph 4, Section 538, P. L. and R., authorized April 1, 1926.

PRINTED IN THE UNITED STATES OF AMERICA

NOTES ON THE HISTORY OF THE USES OF ANALYTICITY IN OPERATOR THEORY

ANGUS E. TAYLOR, University of California, Berkeley

Introduction. Analytic dependence on a complex parameter appears at many places in the study of differential and integral equations. In Fredholm's 1903 paper [15] the solution of the equation

$$f(s) - \lambda \int_a^b k(s, t)f(t)dt = g(s)$$

is presented in the form

$$f(s) = g(s) + \frac{\lambda}{d(\lambda)} \int_a^b D(s, t; \lambda)g(t)dt,$$

provided that $d(\lambda) \neq 0$. Here f and g are members of the function class $C[a, b]$. The function $d(\lambda)$ is an entire analytic function of the complex parameter λ and $D(s, t; \lambda)$, which is a continuous function of (s, t), is also an entire analytic function of λ.

The display of analyticity in the solution of the Fredholm equation of the second kind is an early signal of the important role which analyticity was destined to play in spectral theory. Anticipations of spectral theory itself can be traced at least as far back as 1836, which was the date of publication of the work of Sturm [38], [39] and Liouville [28], [29] on second-order ordinary differential equations and the associated eigenvalue and expansion theorems. In 1904 Hilbert [20] showed how to transform a differential boundary value problem into a problem of solving an integral equation. Moreover, the Green's function of the differential problem turned out, under certain conditions, to be a kernel of Hilbert-Schmidt type, from which followed a great many things of interest.

The utility of analytic functions and the calculus of residues in connection with expansions in series of special functions has been known since the work of Cauchy. In 1894 Poincaré [34] dealt with the problem of the vibrating membrane and related the eigenvalues and eigenfunction expansions to the meromorphic character of the Green's function's dependence on the parameter. In 1908 Birkhoff made explicit the use of the meromorphic behavior of the Green's

Professor Taylor was an undergraduate at Harvard and did his graduate work at Cal. Tech. under A. D. Michal. He stayed on a year at Cal. Tech. spent an NRC Fellowship year at Princeton, then joined UCLA where he has been except for a year as an operations analyst with the USAF in England. Since 1966 he has been Academic Vice-President of the Univ. of Calif. He has spent sabbaticals at Cambridge University, the J. Gutenberg University in Mainz (Fulbright Fellow), and in Geneva and Lausanne. His research has been in functional analysis, and his books are *Calculus* (with G. E. F. Sherwood), *Elementary Differential Equations* (with G. E. F. Sherwood), *Advanced Calculus*, *Calculus with Analytic Geometry*, *Introduction to Functional Analysis*, *General Theory of Functions and Integration*, and *Calculus with Analytic Geometry: Functions of One Variable* (with C. J. A. Halberg, Jr.). *Editor.*

function in dealing with expansion problems associated with ordinary differential equations. See Birkhoff [2], [3]. For general historical background and further references see Dunford and Schwartz [14], pages 1581–1583, and the encyclopedia article of Hilb and Szász [19], pages 1260–1264.

As we now know, from many points of view, the eigenvalue theory and expansion theory of the classical Sturm-Liouville problems for a finite closed interval may be almost completely accounted for as a consequence of the fact that the resolvent operator is meromorphic and has a Mittag-Leffler expansion of a particularly simple sort. (See Taylor [44], pages 342–344.) The relevant formulas are as follows: If T is a symmetric linear operator in Hilbert space such that T^{-1} exists and is compact, then the spectrum $\sigma(T)$ is a countable set $\{\mu_1, \mu_2, \cdots\}$ where $|\mu_n| \to +\infty$, and the resolvent $(\mu - T)^{-1}$ is expressible in the form

$$(\mu - T)^{-1}y = \sum_{k=1}^{\infty} \frac{(y, x_k)}{\mu - \mu_k} x_k$$

for each vector y. Here $\{x_1, x_2, \cdots\}$ is a complete orthonormal set and $Tx_k = \mu_k x_k$.

In the application to Sturm-Liouville theory T is a differential operator such that $Tx = u$ is expressed by

$$-\frac{d}{dt}[p(t)x'(t)] + q(t)x(t) = u(t).$$

The boundary conditions are expressed in the definition of the domain of T. The functions p and q are real and subject to certain conditions.

The pioneering work of F. Riesz. The uses of analyticity in operator theory were originally apparent in a classical context which required no recognition of analyticity in any sense other than what was entailed in speaking of complex-valued analytic functions of a complex variable. This was so because the differential and integral operators under consideration were acting on numerically-valued functions. Thus, for example, in the solution of Fredholm's integral equation, the function $D(s, t; \lambda)$, sometimes called "the first Fredholm minor," is a numerical function of s, t, and λ. It is interesting to examine the progression toward a conscious recognition of the concept of an analytic function with values which are not merely complex numbers but are, instead, elements of a function space or a space of operators, or even an abstract space. A significant step in this direction was made by F. Riesz. The evidences are to be found in his 1913 book [35]. Riesz was studying "linear substitutions" in the theory of systems of linear equations in an infinite number of unknowns. He considered, in particular, completely continuous substitutions acting in the l^2 class of infinite sequences (the classical prototype of a Hilbert space). Riesz asserts that if A is such a substitution and E is the identity substitution, then the inverse substitution $(E - \lambda A)^{-1}$ is a meromorphic function of λ (see page 106 of [35]). We can

see from what Riesz wrote that he was thinking of $(E-\lambda A)^{-1}$ as a function of the complex variable λ with values in the class of bounded linear substitutions on l^2. However, Riesz does not explicitly at this point discuss what it means for such a function to be analytic at a particular point or to have a pole at a particular point. The actual discussion turns on an examination of certain related linear substitutions (and their inverses) when only a finite number of unknowns are considered. These systems are dealt with by using determinants of finite order.

One might suppose that perhaps Riesz meant only to say that the individual terms in the matrix representation of $(E-\lambda A)^{-1}$ are meromorphic functions of λ. That this would be an error is made clear, I think, by what comes a bit further on in Riesz's book. On pages 114–119 he examines $(E-\lambda A)^{-1}$ for the case of an arbitrary bounded linear substitution A (i.e., not merely the completely continuous case). Riesz shows that the class of regular points (those for which $E-\lambda A$ is invertible in an appropriate sense) form an open set in the complex plane and that $(E-\lambda A)^{-1}$ is analytic at each regular point μ in the following sense: $(E-\lambda A)^{-1}$ can be represented as a power series in $\lambda-\mu$ with certain coefficients which are bounded linear substitutions. The series converges in a well-defined sense when λ is sufficiently close to μ. The mode of convergence is that of what we would today call the uniform topology of the operators.

The monograph of Riesz foreshadowed many important later developments in the application of the methods of analytic function theory to operator theory. In particular, Riesz saw the utility of the residue calculus as applied to contour integrals involving $(E-\lambda A)^{-1}$. (See pages 117–121 of [35].)

The move to abstraction. Full realization of the extent to which analyticity could be formulated and applied in *abstract* operator theory was not achieved until twenty-five years later. In a 1938 paper [40] Taylor showed that, if T is a closed linear operator in a complex Banach space and I is the identity operator, the inverse $(\lambda I-T)^{-1}$, if it exists in a suitable sense for at least one λ, is an operator-valued analytic function on the (necessarily open) set of λ's for which $(\lambda I-T)^{-1}$ exists in this suitable sense. This open set of λ's is called the resolvent set of T and $(\lambda I-T)^{-1}$ is then called the resolvent of T. He also showed that if T is bounded and everywhere defined on a Banach space with more than one element, there must be some value of λ for which the resolvent of T is undefined; that is, the spectrum of T, composed of all points not in the resolvent set, is not empty. This result hinges on the use of Liouville's theorem for vector-valued analytic functions (the theorem which asserts that a function which is analytic on the whole complex plane, and also bounded, must be constant in value).

Taylor was not aware, in 1937, of the work of Riesz referred to in the preceding paragraphs. He got his ideas primarily from the works of Norbert Wiener and Marshall Stone. Wiener, in a paper [48] published in 1923, pointed out that Cauchy's integral theorem and much of the classical theory of functions of a complex variable remain valid for functions from the complex plane to a com-

plex complete normed linear space. Wiener did not apply his observations to spectral theory. Taylor read Wiener's paper in 1935 and soon thereafter he read the exposition of spectral theory for operators in Hilbert space in Stone's book [**37**]. Stone established the fact that, if T is a closed linear operator with resolvent $R(\lambda)=(\lambda I-T)^{-1}$, the inner-product $(R(\lambda)f, g)$ (where f and g are vectors in the Hilbert space) is a complex-valued analytic function of λ (page 141 of Stone's book). By further consideration of $(R(\lambda)f, g)$ Stone proved that a bounded and everywhere defined operator T cannot have an empty spectrum. Taylor realized that, by regarding $R(\lambda)$ as an operator-valued function and applying analytic function-theory to it, he could prove in a general Banach space context what Stone had proved in a Hilbert space context.

Soon after this Taylor reflected on the fact that, conceptually, it is not the same thing to say that $R(\lambda)$ is analytic as an operator-valued function as it is to say that $R(\lambda)f$ is, for each f, analytic as a vector-valued function with values in the space from which the vectors f are drawn. This led him (still in 1937) to the discovery (which surprised him very much) of a general theorem transcending spectral theory: *For each λ in an open set D of the complex plane let $A(\lambda)$ be a bounded linear operator from a complex Banach space X to a complex Banach space Y. Suppose that, for each vector x in X, $A(\lambda)x$ is analytic as a function from D to Y. Then $A(\lambda)$ is analytic on D as an operator-valued function, using the uniform topology of operators.*

The converse implication is trivial, but this theorem depends on the deep principle of uniform boundedness. See the proof on page 576 in [**41**]; see also Theorem 4.4-G in [**44**]. Taylor was so surprised by this result that when he announced it at a meeting of the American Mathematical Society in New York on February 26, 1938, he had not thought to consider what might follow from the weaker assumption that $y^*(A(\lambda)x)$ is analytic on D for each x in X and each continuous linear functional y^* on Y. This weaker assumption does indeed imply the analyticity of $A(\lambda)x$ for each x and hence also the further conclusion which Taylor had obtained. This implication—that if a vector-valued function of λ is analytic in the weak topology it is also analytic in the strong topology—had been proved by Dunford (also in 1937) without reference to operator theory. Dunford heard Taylor's presentation in New York early in 1938; until that moment each was unaware of the work of the other on these closely related issues. Dunford's theorem about weak and strong analyticity was published in 1938; see page 354 of [**9**].

Function-theoretic approach to spectral theory. The mere fact that the resolvent $(\lambda I-T)^{-1}$ is an operator-valued analytic function is just the starting point for the uses of analyticity in the spectral theory of the operator T. The open set on which the resolvent is defined is the resolvent set, denoted by $\rho(T)$. The closed set complementary to $\rho(T)$ is the spectrum $\sigma(T)$. Spectral theory of linear operators is the study of linear operators conducted from the point of view of relating the properties of an operator to an analysis of the character of its spectrum and the behavior of the resolvent in the part of $\rho(T)$ near a point

of $\sigma(T)$. If μ is an isolated point of $\sigma(T)$, it is an isolated singular point of the resolvent as an analytic function, and the behavior of the resolvent near μ may be studied (for example) by looking at the Laurent expansion of $(\lambda I-T)^{-1}$ in powers of $\lambda-\mu$. The boundary of $\sigma(T)$ is the natural boundary of the resolvent as an analytic function; that is, $(\lambda I-T)^{-1}$ cannot be continued beyond $\rho(T)$ by analytic continuation.

In its general setting (say in a complex Banach space) spectral theory is the functional analysis counterpart of the theory of eigenvalues of linear transformations in spaces of finite dimension. When the Banach space is finite-dimensional the spectrum of the operator T is a finite set of eigenvalues and $(\lambda I-T)^{-1}$ is a rational function of λ. In the general case the nature of $(\lambda I-T)^{-1}$ can be much more complicated. In particular, the spectrum need not have any isolated points. It may be an uncountably infinite set, and indeed it may fill up a whole two-dimensional area in the plane (e.g., a circular disk). Nevertheless, the analyticity of $(\lambda I-T)^{-1}$ on $\rho(T)$ has far-reaching consequences in spectral theory and in certain cases, at least, much of the spectral theory of a particular operator T may be viewed as the study of both the local and global properties of the function.

Cauchy's formula and operational calculus. An important instrument of spectral theory is the operational calculus based on the definition of an operator $f(T)$ by the "Cauchy formula"

$$f(T) = \frac{1}{2\pi i}\int_C f(\lambda)(\lambda I - T)^{-1}d\lambda,$$

where the integral is extended over a contour suitably located with respect to the spectrum of T and f is a complex function analytic on a neighborhood of the spectrum. The Cauchy formula defines a homomorphism from a ring of functions into the ring of all bounded linear operators acting in the Banach space on which T is defined. For the finite-dimensional case (i.e., for the study of square matrices) there are early anticipations of this operational calculus in the work of Frobenius [16] and others, most explicitly in an 1899 paper by Poincaré [33]. See page 607 of Dunford and Schwartz [13] for more detailed references.

Systematic exploitation of the calculus of residues and of the symbolic operational calculus based on the Cauchy integral formula, as applied to general spectral theory, starts with F. Riesz in his 1913 book and picks up again in the years 1936–1943, with many developments in subsequent years. The work of investigators of normed rings: of Nagumo [32] in 1936, of Mazur [31] in 1938, of Gelfand [17] in 1941, and Lorch [30] in 1943, are part of this general development. The work of Dunford [10], [11], and Taylor [42] in 1943 dealt explicitly with general operator theory. In 1951 Taylor [43] showed how to adapt the Cauchy formula to yield an operational calculus for unbounded closed linear operators. A special case of use of the Cauchy formula to define operators occurs in 1939 work of Hille [21].

Singularities of the resolvent. By studying power series and Laurent series developments of the resolvent $(\lambda I-T)^{-1}$ one comes to some important results in spectral theory. The best-known theorems relate to the C. Neumann expansion of the resolvent:

$$(\lambda I - T)^{-1} = R(\lambda) = \sum_{n=1}^{\infty} \lambda^{-n} T^{n-1},$$

which is valid whenever the series is convergent in the uniform topology. By extension of a theorem of classical function theory (about the radius of convergence of a power series), we have convergence of the Neumann series if $|\lambda| > r(T)$ and divergence if $|\lambda| < r(T)$, where

$$r(T) = \limsup_{n\to\infty} \|T^n\|^{1/n}.$$

As is well known (see Taylor [**44**], pages 262–263), the limit superior here is actually a limit, and $\|T^n\|^{1/n}$ converges to $r(T)$ from above. This was first proved by Gelfand in relation to Banach algebras. The number $r(T)$ is called the spectral radius of T because every point λ of $\sigma(T)$ satisfies $|\lambda| \leqq r(T)$ and some point λ of $\sigma(T)$ is such that $|\lambda| = r(T)$.

A classical theorem of Pringsheim (see, e.g., Landau [**25**], page 72) states that, if a power series with real nonnegative coefficients has the number 1 as its radius of convergence, then the point $z=1$ is a singular point of the function defined by the series. It is possible to extend this theorem to a situation in which the coefficients of the power series lie in a Banach space by assuming that the Banach space has a suitable partial-order structure and that the vector coefficients lie in the positive cone of the space. The theorem, thus generalized, has an application to spectral theory. Without an explanation of all the technicalities this application can be stated as follows: If T is a positive linear operator in an ordered complex Banach space whose positive cone satisfies certain conditions, then the spectral radius of T is a point of the spectrum. This is a generalization of a theorem about matrices known since 1907:

For a square matrix with nonnegative elements the spectral radius is an eigenvalue.

A further result is that, if T is a compact and positive operator in an ordered real Banach space with a positive cone which is total, and if the spectral radius of T is positive, then $r(T)$ is a pole of the resolvent of T. This result was obtained by Krein and Rutman in 1948. For references and an exposition of the function-theoretic methods of proof stemming from Pringsheim's theorem, see pages 261–268 in H. H. Schaefer's book [**36**].

If μ is a point of the spectrum of T and $\{\lambda_n\}$ is a sequence of points in the resolvent set such that $\lambda_n \to \mu$, then $\|(\lambda_n I - T)^{-1}\| \to +\infty$, that is, the norm of the resolvent operator $(\lambda I - T)^{-1}$ becomes infinite as λ approaches a point of the spectrum of T. (See page 260 in Schaefer [**36**].)

An isolated point μ of $\sigma(T)$ is either a pole or an essential singularity of the

resolvent, depending on whether the Laurent development of the resolvent in powers of $\lambda-\mu$ has only a finite or an infinite number of nonvanishing terms in negative powers of $\lambda-\mu$. The coefficient of $(\lambda-\mu)^{-1}$ is a projection (a fact which appears in the pioneering work of F. Riesz). If μ is a pole, it is an eigenvalue of T (see Taylor [**42**], page 660 and [**44**], page 306). If μ is an isolated essential singularity of T, then the *descent* and *defect* of $\mu I-T$ are both infinite. That is, the range of $(\mu I-T)^{n+1}$ is a proper subset of the range of $(\mu I-T)^n$ for $n=1, 2, \cdots$ and the quotient of the entire space by the range of $\mu I-T$ is infinite dimensional. This result is due to D. C. Lay [**27**] (Corollary 2.10). Not much is known about isolated essential singularities of the resolvent.

The boundary of the spectrum is of special interest. If μ is in the boundary of $\sigma(T)$, then either μ is an eigenvalue of T (that is, $\mu I-T$ has no inverse) or the inverse $(\mu I-T)^{-1}$ exists but is discontinuous. (See Gindler and Taylor [**18**], Theorem 3.2.) A boundary point of $\sigma(T)$ need not be an isolated point of the spectrum. However, if μ is a boundary point such that either the defect of $\mu I-T$ is finite or the nullity of $\mu I-T$ is finite and the range of $\mu I-T$ is closed, then μ is a pole of finite rank of the resolvent. This result is essentially contained in Kato's 1958 work [**24**] on perturbation theory, but in this formulation it is due to D. C. Lay ([**27**], Theorem 2.9). For the terminology used here and for a number of related results see Taylor [**42**].

It is natural to seek to classify the points of the spectrum in such a way as to relate the fact that μ belongs to a particular part of the spectrum to certain qualities of the operators $\lambda I-T$ and $(\lambda I-T)^{-1}$ (if the latter exists) when λ is either near or coincident with μ. Studies of such matters may be referred to as studies of the *fine structure* of the spectrum. Taylor and several of his students have made such studies, utilizing various tools: the *state diagrams* (see Taylor [**44**], page 237), the *minimum modulus* of T (see Gindler and Taylor [**18**]), and the indices *ascent*, *descent*, *nullity*, and *defect* (see Taylor [**47**] and Lay [**26**], [**27**]). The very thorough study of fine structure made by Lay is the most closely connected with ideas stemming from the analyticity of the resolvent. It is also very closely related to the work of Kato and others on perturbation theory. Here are two illustrative examples of results in the study of fine structure of the spectrum using the indices ascent, descent, nullity and defect:

If the ascent, descent, nullity, and defect of $\mu I-T$ are all finite, then μ is an isolated point of the spectrum and it is a pole of finite rank of the resolvent.

If the defect and nullity of μ are both infinite, while the ascent and descent are positive and finite, then μ is a pole of infinite rank.

Meromorphic operators. It has long been known (the result goes back to F. Riesz) that, if T is a compact operator, each nonzero point of its spectrum is a pole of the resolvent. This situation suggests consideration of bounded linear operators T of the following sort: the spectrum of T consists of a countable set of points having at most one point of accumulation (which is assumed to be 0, for convenience) and such that each nonzero point of the spectrum is a pole of

the resolvent. For convenience we call such an operator *meromorphic*. If in addition each of the nonzero poles is of finite rank (i.e., if the null space of $\mu I-T$, corresponding to the nonzero pole μ, is finite-dimensional), then the operator is called a Riesz operator. A compact operator is a Riesz operator. Riesz operators T are characterized by the fact that, for each $\lambda\neq 0$, the nullity and defect of $\lambda I-T$ are finite (and necessarily equal). (See Caradus [4], Theorem 2.31.) Lay has given the following simpler characterization in [26]: An operator T is a Riesz operator if and only the defect of $\lambda I-T$ is finite for each $\lambda\neq 0$. Meromorphic operators are characterized by the fact that the descent of $\lambda I-T$ is finite for each $\lambda\neq 0$. This result is due to Lay. Caradus had shown earlier (in [4]) that T is meromorphic if and only if, for each $\lambda\neq 0$, the ascent and descent of $\lambda I-T$ are finite and the range of $(\lambda I-T)^p$ is closed when p is the descent of $\lambda I-T$. For other work on meromorphic operators, Riesz operators, and the closely related notion of a Fredholm operator, see Caradus [5], [6], [7], Kaashoek [23], and Kaashoek and Lay [22].

The resolvent $R(\lambda)$ of a meromorphic operator can be represented globally by a series derived by setting $z=\lambda^{-1}$ and applying the classical Mittag-Leffler expansion process to $R(1/z)$. The resulting series representation of $R(\lambda)$ has been studied by Taylor, Berkson, and Derr (see [45], [46], [1], and [8]). The results from these studies are only fragmentary, and it seems likely that a great deal more can be learned. In order to avoid complications here I shall mention only the simplest results. Let $\lambda_1, \lambda_2, \cdots$ be an enumeration of the nonzero poles of T and let $E_1, E_2, \cdots$ be the corresponding projections. That is, E_n is the residue of $(\lambda I-T)^{-1}$ at λ_n. Suppose the poles are all of the first order and suppose that the series

$$\sum_{n=1}^{\infty}\|\lambda_n E_n\|$$

is convergent. Then, if

$$A=\sum_{n=1}^{\infty}\lambda_n E_n \quad\text{and}\quad B=T-A,$$

A is a meromorphic operator with the same spectrum as T, and B is a quasi-nilpotent operator (one whose spectrum consists of the single point $\lambda=0$) such that $AB=BA=0$. Moreover, the resolvents of A and T are given by the formulas

$$(\lambda I-A)^{-1}=\frac{I}{\lambda}+\sum_{n=1}^{\infty}\left(\frac{1}{\lambda-\lambda_n}-\frac{1}{\lambda}\right)E_n,$$

$$(\lambda I-T)^{-1}=(\lambda I-A)^{-1}+\sum_{n=1}^{\infty}\frac{B^n}{\lambda^{n+1}}\cdot$$

The foregoing situation includes the case in which T is a compact self-adjoint operator in Hilbert space (in which case $T=A$ and $B=0$).

The theory of spectral operators. The theory of spectral operators initiated by Dunford and developed by him and others utilizes analyticity in essential ways. I shall discuss here some aspects of this theory. A bounded operator T in a complex Banach space is called spectral if it possesses a spectral measure, or resolution of the identity, which is countably additive in a suitable sense. The spectral measure is a function $E(\cdot)$ defined on the class of Borel sets in the complex plane; the values of the function are projection operators suitably related to T. The simplest type of spectral operator, the *scalar* type, has an integral representation

$$T = \int \lambda E(d\lambda).$$

In general, a spectral operator has a unique decomposition $T=S+N$, where S is of scalar type while N is quasinilpotent and commutes with T.

The standard decomposition of a spectral operator into its scalar and quasinilpotent parts, together with the representation of the scalar part by means of the spectral measure, is the general counterpart of the canonical Jordan reduction in the finite-dimensional case. In Hilbert space every bounded self-adjoint or normal operator is spectral.

If T is a bounded self-adjoint operator in Hilbert space, its spectrum lies on a finite closed interval of the real axis and the resolvent $(\lambda I-T)^{-1}$ has the property that, if μ is a point of $\sigma(T)$, $\|(\lambda-\mu)(\lambda I-T)^{-1}\|$ is bounded as λ approaches μ along a line perpendicular to the real axis at μ. Furthermore, for each vector x in the Hilbert space, the mapping $\lambda \to (\lambda I-T)^{-1}x$ (which is analytic on the resolvent set $\rho(T)$) has a single-valued maximal analytic continuation into a certain open set containing $\rho(T)$. Let the complement of this open set be denoted by $\sigma(x)$. Now, for any closed set δ of complex numbers, consider the set of all vectors x for which $\sigma(x) \subset \delta$. This set of vectors turns out to be the closed linear manifold which is the range of the projection $E(\delta)$, where $E(\cdot)$ is the resolution of the identity for T.

The foregoing notions can be generalized, and the generalization forms an important part of Dunford's theory of spectral operators. For a detailed survey and many references see Dunford [12]. In the general theory of spectral operators one of the essential requirements on T is that, for each x, the analytic continuation of the mapping $\lambda \to (\lambda I-T)^{-1}x$ give rise to a single-valued maximal analytic continuation. Then the relationship between $E(\delta)$ and the linear manifold of all vectors x such that $\sigma(x) \subset \delta$ is as described earlier. There is a great deal more to the theory, of course. If T is a bounded operator, a sufficient condition for the mapping $\lambda \to (\lambda I-T)^{-1}x$ to have a single-valued maximal analytic continuation for each vector x is that the spectrum of T be nowhere dense in the plane. However, this alone does not suffice to make T spectral, or even to lead to a suitable family of projections from which to construct a spectral measure.

A good deal of attention has been concentrated on situations of the following sort: Suppose the spectrum of T lies in a smooth Jordan curve and that, for each μ in the spectrum, the norm of $(\lambda I-T)^{-1}$ does not grow more rapidly than some negative power of $\lambda-\mu$ as λ approaches μ along a path which is transverse (in a suitable sense) to the Jordan curve at μ. (This is a generalization of the property of the resolvent of a self-adjoint operator in Hilbert space, as referred to earlier). Then the single-valued analytic continuation property is available and it can be shown that a function $E(\cdot)$ exists with the appropriate relation between $E(\delta)$ and the set of x's for which $\sigma(x)\subset\delta$. However, still further requirements are needed to insure that $E(\cdot)$ is countably additive and that T is spectral.

It may be conjectured that it would be fruitful to study further the following general questions: When is a meromorphic operator spectral? Are there generalizations of the Mittag-Leffler expansions of the resolvent which might play a useful role in the study of spectral operators? Are there useful considerations of the indices ascent, descent, nullity, and defect which could enter into the study of spectral operators which are not meromorphic or which have non-discrete spectra?

References

1. E. Berkson, Sequel to a paper of A. E. Taylor, Pacific J. Math., 10 (1960) 767–776.
2. G. D. Birkhoff, Boundary value and expansion problems of ordinary linear differential equations, Trans. Amer. Math. Soc., 9(1908) 219–231.
3. ———, Note on the expansion problems of ordinary linear differential equations, Rend. Circ. Mat. Palermo, 36(1913) 115–126.
4. S. Caradus, Contributions to the theory of operators with meromorphic resolvents, Doctoral dissertation, University of California, Los Angeles, 1965.
5. ———, Operators of Riesz type, Pacific J. Math., 18(1966) 61–71.
6. ———, Operators with finite ascent and descent, Pacific J. Math., 18(1966) 437–449.
7. ———, On meromorphic operators, I., Canadian J. Math., 19(1967) 723–736.
8. J. Derr, and A. E. Taylor, Operators of meromorphic type with multiple poles of the resolvent, Pacific J. Math., 12(1962) 85–111.
9. N. Dunford, Uniformity in linear spaces, Trans. Amer. Math. Soc., 44(1938) 305–356.
10. ———, Spectral theory, Bull. Amer. Math. Soc., 49(1943) 637–651.
11. ———, Spectral theory I. Convergence to projections, Trans. Amer. Math. Soc., 54(1943) 185–217.
12. ———, A survey of the theory of spectral operators, Bull. Amer. Math. Soc., 64(1958) 217–274.
13. N. Dunford and J. T. Schwartz, Linear Operators, Part I: General Theory. Interscience, New York, 1958.
14. ——— and ———, Linear Operators, Part II: Spectral Theory. Interscience, New York, 1963.
15. I. Fredholm, Sur une classe d'équations fonctionnelles, Acta Math., 27(1903) 365–390.
16. G. Frobenius, Über die cogredienten Transformationen der bilinearen Formen, S.-B. der K. Preuss. Akad. Wiss. Berlin, (1896) 7–16.
17. I. M. Gelfand, Normierte Ringe, Matem. Sbornik, 9(51) (1941), 3–24.
18. H. A. Gindler and A. E. Taylor, The minimum modulus of a linear operator and its use in spectral theory, Studia Math., 22(1962) 15–41.
19. E. Hilb and O. Szász, Allgemeine Reihenentwicklungen, Encyklopädie der Math. Wiss., II C II(1922) 1229–1276.

20. D. Hilbert, Grundzüge einer allgemeinen Theorie der linearen Integralgleichungen, Wiss. Nachrichten Akad. Wiss. Göttingen, Math.-Phys. Kl. (1904) 49–91.

21. E. Hille, Notes on linear transformations II. Analyticity of semigroups, Ann. of Math. (2) 40(1939) 1–47.

22. M. A. Kaashoek and D. C. Lay, On operators whose Fredholm set is the complex plane, Pacific J. Math., 21(1967) 275–278.

23. M. A. Kaashoek, On the Riesz set of a linear operator, Nederl. Akad. Wetensch., Proc. Ser. A, 71(1968) 46–53.

24. T. Kato, Perturbation theory for nullity, deficiency, and other quantities of linear operators, J. Analyse Math., 11(1958) 261–322.

25. E. Landau, Darstellung und Begründung einiger neuerer Ergebnisse der Funktionentheorie, Chelsea, New York, 1946.

26. D. C. Lay, Studies in spectral theory using ascent, descent, nullity and defect, Doctoral dissertation, University of California, Los Angeles, 1966.

27. ———, Spectral analysis using ascent, descent, nullity and defect (to be published in Math. Ann.).

28. J. Liouville, Sur le développement des fonctions en séries dont les divers termes sont assujétis à satisfaire à une même équation différentielle du second ordre contenant un paramètre variable, J. Math. Pures Appl., (1) 1(1836) 253–265.

29. ———, Démonstration d'un théorème dû à M. Sturm et relatif à une classe de fonctions transcendantes, *ibid.* (1) 1(1836) 269–277.

30. E. R. Lorch, The theory of analytic functions in normed abelian vector rings, Trans. Amer. Math. Soc., 54(1943) 414–425.

31. S. Mazur, Sur les anneaux linéaires, C.R. Acad. Sci. Paris, 207(1938) 1025–1027.

32. M. Nagumo, Einige analytische Untersuchungen in linearen metrischen Ringen, Japanese J. Math., 13(1936) 61–80.

33. H. Poincaré, Sur les groupes continus, Trans. Camb. Phil. Soc., 18(1899) 220–255.

34. ———, Sur les équations de la physique mathématique, Rend. Circ. Mat. Palermo, 8(1894) 57–155.

35. F. Riesz, Les systèmes d'équations linéaires à une infinité d'inconnues, Gauthier-Villars, Paris, 1913.

36. H. H. Schaefer, Topological Vector Spaces, Macmillan, New York, 1966.

37. M. H. Stone, Linear Transformations in Hilbert Space, Amer. Math. Soc., New York, 1932.

38. C. Sturm, Sur les équations différentielles du second ordre, J. Math. Pures Appl., (1) 1(1836) 106–186.

39. ———, Sur une classe d'équations à différences partielles, *ibid.* (1) 1(1836) 373–444.

40. A. E. Taylor, The resolvent of a closed transformation, Bull. Amer. Math. Soc., 44(1938) 70–74.

41. ———, Linear operations which depend analytically on a parameter, Ann. of Math., 39(1938) 574–593.

42. ———, Analysis in complex Banach spaces, Bull. Amer. Math. Soc., 49(1943) 652–669.

43. ———, Spectral theory of closed distributive operators, Acta Math., 84(1951) 189–224.

44. ———, Introduction to Functional Analysis, Wiley, New York, 1958.

45. ———, Mittag-Leffler expansions and spectral theory, Pacific J. Math., 10(1960) 1049–1066.

46. ———, Spectral theory and Mittag-Leffler type expansions of the resolvent, Proc. International Symposium on Linear Spaces, The Hebrew University, Jerusalem (1960) 426–440.

47. ———, Theorems on ascent, descent, nullity, and defect of linear operators, Math. Ann., 163(1966) 18–49.

48. N. Wiener, Note on a paper of M. Banach, Fund. Math., 4(1923) 136–143.

Supplementary References

At the suggestion of the referee some additional general references have been added. The most relevant pages in each of the cited works are indicated in parentheses after the citation.

J. Dieudonné, Foundations of Modern Analysis, Academic Press, New York, 1960 (pp. 308–311; 319–323; 329–351).

T. Kato, Perturbation Theory for Linear Operators, Springer Verlag, Berlin and Heidelberg, 1966 (pp. 364–426).

M. A. Naimark, Normed Rings, Nordhoff, Groningen, 1959 (pp. 65–66; 82–83; 173–175; 192–195; 202–207).

K. Yosida, Functional Analysis, Springer Verlag, Berlin and Heidelberg, 1965 (pp. 209–212; 225–231).

A. C. Zaanen, Linear Analysis, North Holland Publishing Co., Amsterdam, 1953 (pp. 344–350; 462–495; 500–509).

DENSITY QUESTIONS IN ALGEBRAIC NUMBER THEORY

L. J. GOLDSTEIN, University of Maryland

Very often, the number theorist bases conjectures on empirical investigations. Even before the invention of the electronic computer, number theorists spent much time doing calculations, the results of which suggested possibly true statements. After the empirical stage of his investigation is completed, the number theorist then tries to supply proofs for his conjectures. It is here where the number theorist applies a formidable armada of high-powered machinery, ranging from analytic function theory to algebraic geometry. It is most surprising that even the most innocently conceived conjecture may lead into a vast jungle of very difficult and technical mathematics. But such is the nature of number theory. In this lecture, I should like to discuss a set of conjectures which typify the process of number-theoretic creation as we have described it: These conjectures originate out of empirical investigation and those few that we are able to prove seem to lead us far afield for their proofs.

1. Gauss' conjecture. Let us denote by $\mathbf{Z}$ the rational integers, p an odd prime, a an arbitrary integer, and $Z_p^\times$ the group of nonzero residue classes mod p. Since $Z_p^\times$ is the multiplicative group of a finite field, a well-known result asserts that $Z_p^\times$ is cyclic of order $p-1$. We say that a is a *primitive root* modulo p if $(a, p)=1$ and if its residue class $\bar{a}$ in $Z_p^\times$ is a generator of $Z_p^\times$.

LEMMA 1.1. *The number a is a primitive root modulo p if and only if $(a, p)=1$ and $a^\nu \not\equiv 1 \pmod p$ for $\nu=1, 2, \cdots, p-1$.*

Larry Goldstein received his Princeton PhD under G. Shimura in 1967. He was a Gibbs lecturer at Yale for two years before his present associate professorship at Maryland. His main research is in analytic and algebraic number theory, and his book, *Analytic Number Theory*, is scheduled to appear. *Editor.*

Note that by Fermat's Little Theorem, if $(a, p)=1$, then

$$a^{p-1} \equiv 1 \pmod{p}.$$

From now on, let us fix a, and let us define

$$\mathcal{A}(a) = \{p \mid p \text{ is prime and } a \text{ is a primitive root modulo } p\}.$$

It may be that $\mathcal{A}(a)$ is empty. For example, if a is a perfect square, say x^2, with $(p, a)=1$, then

$$x^{p-1} \equiv 1 \pmod{p}$$

by Fermat's Little Theorem, so that

$$a^{(p-1)/2} \equiv 1 \pmod{p}.$$

Therefore, if $p \nmid x$, p odd, then $p \notin \mathcal{A}(a)$. However, if $p \mid x$, then it is certainly true that $p \notin \mathcal{A}(a)$. Therefore, we have shown that $\mathcal{A}(a)=\emptyset$ if a is a perfect square. Moreover, since $(-1)^2=+1$, we see that $p \notin \mathcal{A}(-1)$ if $p-1>2$. Therefore, since -1 is a primitive root modulo 3, we have proved that $\mathcal{A}(-1)=\{3\}$.

By means of laborious calculations, Gauss investigated the case $a=10$ and arrived at the following conjecture, which is stated in Article 303 of his *Disquisitiones Arithmeticae*:

CONJECTURE A: $\mathcal{A}(10)$ *is infinite.*

In the next sections, we shall present some heuristic evidence for this conjecture, as well as some more general conjectures which seem to be true.

2. Artin's conjecture. In a conversation with Hasse in 1927, Artin made the following conjecture:

CONJECTURE B: *Suppose that a is not -1 and not a perfect square. Then $\mathcal{A}(a)$ is infinite.*

This conjecture was not just a wild guess, but followed from a very compelling probabilistic argument which Artin advanced. In order to trace Artin's line of thought, we must first define a few notions.

Let $\mathcal{S}$ be a set of primes (finite or infinite), and let x be a positive real number. Let $\pi(x)$ denote the number of primes $\leqq x$, and let $\pi(x, \mathcal{S})$ denote the number of primes in $\mathcal{S}$ which are $\leqq x$. We say that $\mathcal{S}$ has a *(natural) density* if

$$\lim_{x\to\infty} \pi(x, \mathcal{S})/\pi(x)$$

exists. The value of the limit is called the *density of* $\mathcal{S}$ and is denoted $d(\mathcal{S})$. We clearly have

$$0 \leqq d(\mathcal{S}) \leqq 1.$$

Moreover, if $d(\mathcal{S})>0$, then $\mathcal{S}$ is infinite since $\pi(x)\to\infty$ as $x\to\infty$. In a few moments

we shall reformulate Conjecture B in the form of a statement about densities. But first we must state some preliminary information about algebraic number theory.

Let $\boldsymbol{K}$ be an algebraic number field, that is, a finite, algebraic extension of $\boldsymbol{Q}$. Let $\mathfrak{O}$ be the ring of integers of $\boldsymbol{K}$, that is, the integral closure of $\boldsymbol{Z}$ in $\boldsymbol{K}$. If p is an ordinary prime, then $p\mathfrak{O}$ is an ideal of $\mathfrak{O}$, but is usually no longer a prime ideal. However, $p\mathfrak{O}$ can be written as a product of powers of prime ideals of $\mathfrak{O}$ (since $\mathfrak{O}$ is a Dedekind domain):

$$p\mathfrak{O} = \mathfrak{P}_1^{e_1} \cdots \mathfrak{P}_g^{e_g}.$$

It is a general fact from algebraic number theory that $g \leqq \deg(\boldsymbol{K}/\boldsymbol{Q})$. We say that p *splits completely* in $\boldsymbol{K}$ if $g = \deg(\boldsymbol{K}/\boldsymbol{Q})$. Here is a basic theorem which one meets in the analytical portion of algebraic number theory.

THEOREM 2.1 (Dirichlet). *Let $n = \deg(\boldsymbol{K}/\boldsymbol{Q})$ and let $\mathcal{S}$ denote the set of all primes which split completely in $\boldsymbol{K}$. Then $\mathcal{S}$ has a density and*

$$d(\mathcal{S}) = 1/n.$$

Let q be a prime and let $\boldsymbol{L}_q$ denote the splitting field over $\boldsymbol{Q}$ of the polynomial $X^q - a$. We get $\boldsymbol{L}_q$ from $\boldsymbol{Q}$ in two steps. First we adjoin to $\boldsymbol{Q}$ a primitive qth root of unity ζ_q. Then we adjoin to $\boldsymbol{Q}(\zeta_q)$ any qth root of a, say the real value of $a^{1/q}$. Then,

$$\boldsymbol{L}_q = \boldsymbol{Q}(\zeta_q, a^{1/q}). \tag{1}$$

$\boldsymbol{L}_q/\boldsymbol{Q}(\zeta_q)$ is a Galois extension of degree either 1 or q with cyclic Galois group. (The extension is a so-called Kummer extension.) Also, $\boldsymbol{Q}(\zeta_q)/\boldsymbol{Q}$ is a Galois extension of degree $q-1$ with cyclic Galois group. Thus, $\boldsymbol{L}_q/\boldsymbol{Q}$ is a Galois extension with solvable Galois group and

$$\deg(\boldsymbol{L}_q/\boldsymbol{Q}) = q - 1 \quad \text{or} \quad q(q-1), \tag{2}$$

depending on the value of a.

From the tool box of the algebraic number theorist, we quote the following result:

THEOREM 2.2. *p splits completely in $\boldsymbol{L}_q \Leftrightarrow p \equiv 1 \pmod{q}$ and*

$$a^{(p-1)/q} \equiv 1 \pmod{q}.$$

Combined with Lemma 1.1, this yields the following:

THEOREM 2.3. *a is a primitive root modulo p if and only if for each prime q, the prime p does not split completely in $\boldsymbol{L}_q$.*

For $\boldsymbol{K}$ an algebraic number field, let $\mathrm{Spl}(\boldsymbol{K})$ denote the set of all primes which split completely in $\boldsymbol{K}$; let $\mathcal{P}$ denote the set of all primes. Then, by Theorem 2.3, we can assert the following:

COROLLARY 2.4. $\mathfrak{A}(a)=\bigcap_q (\mathcal{P}-\mathrm{Spl}(L_q))$.

Now for Artin's probabilistic argument: By Dirichlet's theorem, $\mathrm{Spl}(L_q)$ has a density and $d(\mathrm{Spl}(L_q))=1/\deg(L_q/Q)$. Therefore, $\mathcal{P}-\mathrm{Spl}(L_q)$ has a density and

$$d(\mathcal{P} - \mathrm{Spl}(L_q)) = 1 - 1/\deg(L_q/Q).$$

Therefore, from Corollary 2.4, we might guess that $\mathfrak{A}(a)$ has a density and that

$$(*) \qquad d(\mathfrak{A}(a)) = \prod\nolimits_q (1 - 1/n(q)), \qquad n(q) = \deg(L_q/Q).$$

Let us see how (*) fits in with Conjecture B. First of all, it is easy to check that if $a\neq -1$ then $n(q)=q(q-1)$ for all but a finite number of q. Therefore, the product converges for $a\neq -1$. For $a=-1$, the product diverges to 0. Thus, if $a\neq -1$, the product can converge to 0 if and only if one of the factors $=0$, and this in turn if and only if $n(q)=1$ for some q. But it is trivial to check that $n(q)\geq q-1>1$ if $q>2$. Therefore, the product $=0$ if and only if $n(2)=1$. But $L_2=Q(a^{1/2})$, so that $n(2)=1$ if and only if a is a perfect square. Therefore, we conclude that if $a\neq -1$ and $a\neq b^2$, then the product is positive, so that $d(\mathfrak{A}(a))>0$, which implies Conjecture B.

Thus, the heuristic arguments of Artin seemed to fit the facts such as they were known at the time. However, experimental calculations by D. H. Lehmer cast a serious doubt as to whether the true value of the density of $\mathfrak{A}(a)$ was given by (*). In the face of this disagreement between conjecture and evidence, it was necessary to reexamine the reasoning which led to (*). Let us consider the probabilistic event "a randomly chosen prime belongs to $\mathcal{P}-\mathrm{Spl}(L_q)$." Dirichlet's Theorem may be interpreted as saying that the probability of this event is $1/n(q)$. We then get the probability that a randomly chosen prime belongs to the intersection of all $\mathcal{P}-\mathrm{Spl}(L_q)$ by multiplying the corresponding probabilities. This is valid, as every student of probability knows, only when the events are pairwise independent. Therefore, what probably goes wrong is that something analogous to probabilistic independence is violated. Of course, all of our analogies with probability theory are only of heuristic value. But they seem to lead somewhere in this case! For, upon close inspection, we see that the fields L_q are not "independent" of one another, that is, it is not true that $L_q\cap L_{q'}=Q$ for $q\neq q'$. Therefore, if we wish to make a statement like (*), it is necessary to somehow take into account this dependence.

By Corollary 2.4,

$$\mathfrak{A}(a) = \mathcal{P} - \bigcup_q \mathrm{Spl}(L_q).$$

Note, however, that the primes which split completely in two fields L_{q_1} and L_{q_2} are subtracted twice on the right hand side of (3). In an attempt to count each prime in $\mathfrak{A}(a)$ once and only once, let us add back in those primes which were removed twice to get

$$\mathfrak{A}(a) = \mathcal{P} - \bigcup_q \mathrm{Spl}(L_q) + \bigcup_{\substack{q_1,q_2\\ q_1\neq q_2}} \mathrm{Spl}(L_{q_1}) \cap \mathrm{Spl}(L_{q_2}).$$

In adding the last term, however, we have counted twice the primes which split completely in three fields L_{q_1}, L_{q_2}, L_{q_3}. Therefore, let us correct this by writing

$$\begin{aligned}\mathcal{A}(a) = \mathcal{P} - \bigcup_{q} \mathrm{Spl}(L_q) + \bigcup_{\substack{q_1, q_2 \\ q_1 \neq q_2}} \mathrm{Spl}(L_{q_1}) \cap \mathrm{Spl}(L_{q_2}) \\ - \bigcup_{\substack{q_1, q_2, q_3 \\ q_i \text{ distinct}}} \mathrm{Spl}(L_{q_1}) \cap \mathrm{Spl}(L_{q_2}) \cap \mathrm{Spl}(L_{q_3}).\end{aligned}$$

But now the primes which split completely in four fields have been subtracted twice, so we must add them back in, and the process continues. Eventually, we arrive at a formula for $\mathcal{A}(a)$ in which each prime is counted exactly once. If $q_1, q_2, \cdots, q_r$ are distinct primes, $k = q_1 \cdots q_r$, let us define L_k to be the composite

$$L_k = L_{q_1} \cdot \cdots \cdot L_{q_r}.$$

Then

$$\mathrm{Spl}(L_{q_1}) \cap \cdots \cap \mathrm{Spl}(L_{q_r}) = \mathrm{Spl}(L_k).$$

Therefore, we may write our formula for $\mathcal{A}(a)$ in the form

$$\begin{aligned}\mathcal{A}(a) = \mathcal{P} - \bigcup_{q} \mathrm{Spl}(L_q) + \bigcup_{\substack{q_1, q_2 \\ q_i \text{ distinct}}} \mathrm{Spl}(L_{q_1 q_2}) \\ - \bigcup_{\substack{q_1, q_2, q_3 \\ q_i \text{ distinct}}} \mathrm{Spl}(L_{q_1 q_2 q_3}) + \cdots.\end{aligned} \tag{3}$$

We have defined L_k for each positive, square-free integer. Let $n(k) = \deg(L_k/\boldsymbol{Q})$. Then by Dirichlet's Theorem and (3), we can conjecture that

$$d(\mathcal{A}(a)) = 1 - \sum_{q} n(q)^{-1} + \sum_{\substack{q_1, q_2 \\ q_i \text{ distinct}}} n(q_1 q_2)^{-1} - \cdots. \tag{4}$$

By rewriting the right hand side of (4), we derive the following conjecture:

CONJECTURE C: *$\mathcal{A}(a)$ has a natural density, and*

$$d(\mathcal{A}(a)) = \sum_{k} \mu(k)/n(k), \qquad n(1) = 1,$$

where $\mu(k)$ denotes the Möbius function and the sum runs over all positive square-free integers k (including 1).

It is Conjecture C that agrees with the experimental evidence. Note, however, that from the form of the sum in Conjecture C, it is no longer evident that $d(\mathcal{A}(a)) > 0$ if $a \neq -1$ and $a \neq b^2$. Also, it must be checked that the series converges. Both points are answered by the following theorem.

THEOREM 2.5 (Hooley [3]). *Let k be a positive square-free integer, let h denote the largest positive integer such that a is an h-th power, and let*

$$k_1 = k/(h, k),$$
$$a_1 = \text{the square-free part of } a,$$
$$\epsilon(k) = \begin{cases} 2 & \text{if } k \text{ is divisible by } 2a_1 \text{ and } a_1 \equiv 1 \pmod 4 \\ 1 & \text{otherwise.} \end{cases}$$

Then $n(k) = k_1\phi(k)/\epsilon(k)$, where $\phi(k)$ denotes Euler's function.

As an immediate consequence of Hooley's theorem, we deduce two corollaries.

COROLLARY 2.6. *The sum $\sum_k \mu(k)/n(k)$ converges absolutely.*

COROLLARY 2.7. *If k and a are relatively prime, then*

$$n(k) = k\phi(k).$$

Using Hooley's theorem, we can write the sum of Conjecture C as a product, so that we may revise Conjecture C as follows:

CONJECTURE D: *$\mathfrak{A}(a)$ has a natural density and*

$$d(\mathfrak{A}(a)) = \begin{cases} C(k), & a_1 \not\equiv 1 \pmod 4) \\ C(k)\cdot\left[1 - \mu(|a_1|) \prod\limits_{\substack{q\mid h \\ q\mid a_1}} (q-2)^{-1} \prod\limits_{\substack{q\nmid h \\ q\mid a_1}} (q^2-q-1)^{-1}\right], & a_1 \equiv 1 \pmod 4), \end{cases}$$

where

$$C(k) = \prod_{q\mid h} (1 - (q-1)^{-1}) \prod_{q\nmid h} (1 - \phi(q^2)^{-1}).$$

This is our final form of Artin's conjecture. Implicit in the statement of Conjecture D is the statement that if $a \neq -1$ and a is not a perfect square, then $d(\mathfrak{A}(a)) > 0$. For then $|a_1| \neq 1$. Since $C(k) > 0$, we see that (mod Conjecture D)

$$d(\mathfrak{A}(a)) = 0 \Leftrightarrow a_1 \equiv 1 \pmod 4 \quad \text{and} \quad \mu(|a_1|) = 1$$

and

$$\prod_{\substack{q\mid h \\ q\mid a_1}} (q-2)^{-1} \prod_{\substack{q\nmid h \\ q\mid a_1}} (q^2-q-1)^{-1} = 1.$$

The last of the three conditions on the right can be satisfied only when $|a_1| = 1$, 2, or 3. But of these three possibilities only $|a_1| = 1$ is consistent with the remaining two conditions. Therefore Conjecture D implies

$$d(\mathfrak{A}(a)) = 0 \Leftrightarrow |a_1| = 1 \qquad (5)$$
$$\Leftrightarrow a = -1 \quad \text{or} \quad a = b^2.$$

3. Bilharz's Theorem. Let k be a finite field with q elements, $k[t]$ the ring of polynomials over k in an indeterminate t, and $K=k(t)$ the field of rational functions in t with coefficients in K. The field K is the simplest example of an algebraic function in one variable. The arithmetic properties of such fields parallel the arithmetic of Q, with $k[t]$ playing the role of the rational integers. In many ways, the arithmetic of K is even simpler than that of Z, so that often number theorists use function fields as a testing ground for conjectures about the rational integers. This testing process consists of reformulating a problem about Q or Z into an analogous problem about K or $k[t]$, respectively, and then solving the analogous problem.

In 1935, Bilharz [1], a student of Hasse, formulated and proved the analogue of Artin's conjecture. The role of the rational primes is played by the monic, irreducible polynomials $P \in k[t]$. If P is such a polynomial, then the *norm of P*, denoted NP, is defined by

$$NP = q^r, \qquad r = \deg(P).$$

The quotient ring

$$K_P = k[t]/Pk[t], \qquad P \text{ monic, irreducible},$$

is a finite field with NP elements. The multiplicative group $K_P^\times$ of K_P is cyclic. Suppose that $A \in K$ is not divisible by P. We say that A is a *primitive root modulo P* if A mod $Pk[t]$ generates $K_P^\times$. Given $A \in K$, we can define

$$\mathcal{A}(A) = \{P \mid A \text{ is a primitive root modulo } P\}.$$

It is easy to check that if A is an *r-th* power for some r dividing $q-1$, then $\mathcal{A}(A)=\emptyset$. In analogy with the situation in Q, we can formulate a conjecture.

CONJECTURE A′: *If A is not an r-th power for any prime r dividing $q-1$, then $\mathcal{A}(A)$ is infinite.*

Conjecture A' was proved in the cited work of Bilharz. The most interesting feature of Bilharz's paper is that he proves Conjecture A′ only by assuming a deep result, at the time conjectured but not proved, known as the "Riemann hypothesis for function fields over finite fields." The conjecture was settled by André Weil in 1941 [4], so that the gap in Bilharz's argument was filled.

Let $\mathcal{S}_0$ be the set of all monic, irreducible polynomials in $k[t]$, and let $x \geqq 0$. For $\mathcal{S} \subseteq \mathcal{S}_0$, define

$$\pi_K(x, \mathcal{S}) = \sum_{\substack{P \in \mathcal{S} \\ NP \leqq x}} 1, \qquad \pi_K(x) = \pi_K(x, \mathcal{S}_0).$$

We say that $\mathcal{S}$ has a *natural density* if

$$\lim_{x \to \infty} \frac{\pi_K(x, \mathcal{S})}{\pi_K(x)}$$

exists. We can formulate the analogues of the density conjectures in the function field case. However, the situation here is very much different from the preceding case. The set $\mathfrak{a}(A)$ usually does not have a natural density. However, it is possible to define a new concept of density (Dirichlet density) with respect to which the analogues of the density conjectures are true. The proofs of these results are contained in Bilharz's paper.

4. Hooley's Theorem. Let L be an algebraic number field. If $\mathfrak{A}$ is an ideal of the ring of integers $\mathfrak{O}_L$, the *norm of* $\mathfrak{A}$ denoted $N\mathfrak{A}$, is the number of elements in the (finite) ring $\mathfrak{O}_L/\mathfrak{A}$. The *Dedekind zeta function* of L is defined by

$$\zeta_L(s) = \sum_{\mathfrak{A}} N\mathfrak{A}^{-s},$$

where $\mathfrak{A}$ runs over all ideals of $\mathfrak{O}_K$ and s is a complex variable. The series on the right converges absolutely for $\mathrm{Re}(s) > 1$. Moreover, for s in this half-plane,

$$\zeta_L(s) = \prod_{\mathfrak{p}} (1 - N\mathfrak{p}^{-s})^{-1}, \tag{6}$$

where $\mathfrak{p}$ runs over all prime ideals of $\mathfrak{O}_L$. The product of (6) converges absolutely for $\mathrm{Re}(s) > 1$. Therefore,

$$\zeta_L(s) \neq 0 \qquad (\mathrm{Re}(s) > 1). \tag{7}$$

It is possible to show that $\zeta_L(s)$ can be analytically continued to a meromorphic function on the whole s-plane. The continued function (also denoted $\zeta_L(s)$) has only one pole, a simple pole at $s = 1$ with residue 1. Moreover, $\zeta_L(s)$ satisfies a functional equation connecting its behavior at s with its behavior at $1 - s$. One consequence of this functional equation is that the zeros of $\zeta_L(s)$ in the half-plane $\mathrm{Re}(s) < 0$ are known. These zeros are called *trivial zeros*. By (7), all nontrivial zeros of $\zeta_L(s)$ lie in the strip

$$0 \leqq \mathrm{Re}(s) \leqq 1.$$

There is strong evidence in favor of the following conjecture.

CONJECTURE (Riemann Hypothesis): *All nontrivial zeros of* $\zeta_L(s)$ *lie on the line* $\mathrm{Re}(s) = 1/2$.

The special case $L = Q$ of this celebrated conjecture was first stated by Riemann in 1860. Although the Riemann hypothesis has received the attention of many of the greatest mathematicians of the last 100 years, it remains unproved, and is one of the most significant unsolved problems of contemporary mathematics.

There is a link between the Riemann hypothesis and Conjecture C (the most general form of Artin's conjecture)—namely, Hooley [3], has proved the analogue of Bilharz's theorem:

THEOREM 4.1. *Assume that the Riemann hypothesis is true for each of the fields* L_k. *Then Conjecture* C *is true.*

5. Analogues of Artin's conjecture. It is possible to generalize the heuristic argument which gave rise to Conjecture C: Suppose that $\mathcal{S}$ is a set of rational primes, and suppose that for each $q \in \mathcal{S}$ there is given a number field L_q. Let $\mathcal{A} = \mathcal{A}(\mathcal{S}, \{L_q\})$ denote the set of rational primes which do not split completely in each L_q for $q \in \mathcal{S}$. Let us make a conjecture about the natural density of $\mathcal{A}$.

For $k = q_1 \cdots q_r$, $q_i \in \mathcal{S}$, set

$$L_k = L_{q_1} \cdots L_{q_r},$$
$$n(k) = \deg(L_k/\boldsymbol{Q}).$$

Define $L_1 = \boldsymbol{Q}$, so that $n(1) = 1$. Using the same arguments as in Paragraph 2, we can formulate another conjecture.

CONJECTURE E: *Suppose that*

$$\sum_k n(k)^{-1}$$

converges, where the sum runs over all k for which $n(k)$ is defined. Then $\mathcal{A}$ has a natural density

$$d(\mathcal{A}) = \sum_k \mu(k) n(k)^{-1}.$$

Conjecture E clearly contains Conjecture C as a special case, namely for $\mathcal{S} = \{\text{all rational primes}\}$, $L_q = \boldsymbol{Q}(\zeta_q, a^{1/q})$ $(q \in \mathcal{S})$. There are only two special cases for which Conjecture E has been verified. When $\mathcal{S}$ is finite, Conjecture E can be easily checked using Dirichlet's theorem. When $\mathcal{S}$ is infinite, however, Conjecture E is very difficult. The only case known is now given.

THEOREM 5.1 (Goldstein [2]). *Suppose that*

$$L_q \supseteq \boldsymbol{Q}(\zeta_{q^2})$$

holds for all but a finite number of $q \in \mathcal{S}$. Then Conjecture E *is true. In particular, Conjecture* E *is true if $\mathcal{S} = \{$all rational primes$\}$ and*

$$L_q = \boldsymbol{Q}(\zeta_{q^2}, a^{1/q}) \qquad (a \in \boldsymbol{Z},\ q \in \mathcal{S}).$$

Theorem 5.1 is tantalizingly close to Artin's conjecture. One might hope that the methods used to prove Theorem 5.1 could be appropriately generalized to prove Artin's conjecture. However, it appears that Conjecture E is of a much higher order of difficulty and any hopes in that direction are overly optimistic.

6. Conclusion. In this talk I have tried to indicate how a number theorist comes by his conjectures. In some sense, the combination of intuition, deduction, and heuristic arguments by means of which we have arrived at our conjectures, is a typical way in which many mathematicians work. There is much that we have been forced to omit. For example, it is possible to formulate Con-

jecture E as a conjecture about Haar measure on a certain compact topological group. In this formulation Conjecture E can be thought of as a generalization of Dirichlet's theorem to infinite-dimensional extensions L of Q. For an exposition of this theory, the reader is referred to [2]. If I have said little about methods of proof, it is because there are only a few theorems now proved in the subject. I hope that this talk will generate enough interest to remedy this appalling situation.

This paper is the text of an invited address delivered by the author under the title "On a Conjecture of Artin" at the Northeastern Sectional Meeting on June 20, 1969.

Research supported by NSF Grant GP-13872.

References

1. H. Bilharz, Primdivisoren mit vorgegebener Primitivwürzel, Math. Ann., 114 (1937) 476–492.
2. L. Goldstein, Analogues of Artin's conjecture, Trans. Amer. Math. Soc., (to appear).
3. C. Hooley, On Artin's conjecture, J. Reine Angew. Math., 225 (1967) 209–220.
4. A. Weil, Sur les Courbes Algébriques et les Variétés qui s'en Déduisent, Hermann, Paris, 1948.

MUSIMATICS or THE NUN'S FIDDLE*

A. L. LEIGH SILVER, Fellow of the Institute of Musical Instrument Technology, England

1. The divine ratio. "*Abominandum!*" said Cicero as he went a purler over a hidden obstacle—"*quid est quod?*"—and scrabbling in the undergrowth he uncovered an ancient monument. The lettering was illegible but the design—a cylinder circumscribing a sphere—was clearly that which Archimedes, who was killed in the fall of Syracuse 212 B.C., had charged his friends to inscribe on his tombstone. Since Cicero made this discovery about 75 B.C., the tomb has again been lost, probably forever.

Archimedes transformed empirical knowledge into theoretical science and developed the integral calculus which he said would be used by mathematicians "as yet unborn." In keeping with Aristotle's dictum that "it is proper to consider the similar even in things far distant from each other," he considered it highly significant that the cylinder and inscribed sphere, as regards surface

* A symbolic title with Chaucerian overtones. This one-stringed instrument, better known as the 'Marine Trumpet', has clarion qualities well suited for trans-Atlantic communication.

A. L. Leigh Silver writes that he is a 3M man: medicine, music, and maths. Son of a professional organist, he is an Oxford and London educated physician and presently is employed by the 7520 U.S.A.F. Hospital. He is a fellow of the British Medical Association, Fellow of the Inst. of Musical Instrument Technology, and Hon. Fellow Mercator Music Foundation. *Editor.*

area and volume, are in the ratio 3:2 and that the same relationship exists between the frequencies of an important musical interval.

The ear is very sensitive to this interval—the perfect fifth—and it has been used for tuning instruments from the earliest times. A power of 3/2 can never equal a power of 2/1 and superimposed perfect fifths will never arrive at an octave duplication of the fundamental. On a keyboard instrument, however, we find that twelve fifths pass through the twelve semitones of the chromatic scale and finish on the seventh octave of the fundamental note (Figure 1), which means that the fifths are not all perfect and somewhere the difference between 12 perfect fifths and 7 octaves has been lost. This difference $(3/2)^{12} \div 2^7 = 3^{12}/2^{19}$ is called a 'Pythagorean comma.'

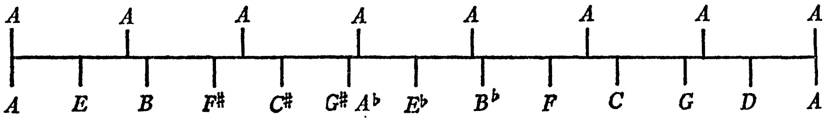


FIG. 1

For tuning purposes the series of fifths is kept within the limits of an octave by descending an octave each time this limit is exceeded (dotted lines Figure 2).

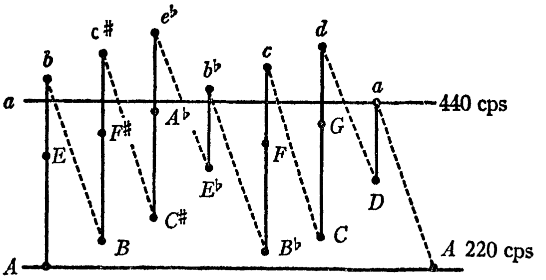


FIG. 2

The ancient Greeks and Chinese calculated the Pythagorean comma which equals about 24 cents (the **cent** being one twelve hundredth part of an octave, or the base two logarithm of the ratio multiplied by 1200).

About 40 B.C., King-Fang, a scholar of the Han dynasty, continued the series of superimposed fifths in order to find a closer approximation to an octave. His first improvement came with the 41st fifth which was less than 24 octaves by about 20 cents. Not content, he carried on until he came to the 53rd fifth, which exceeds 31 octaves by about 3.6 cents. This excellent approximation was later recommended by Mercator and the 53 note octave was incorporated in several instruments including Bosanquet's *Enharmonic Harmonium* which was exhibited in the South Kensington Museum in 1876. In the hope of achieving immortality, I carried on Fang's calculations (without an abacus) and found the next better approximations to be:

$$(3/2)^{306} < 2^{179} \text{ by about } 1.8 \text{ cents}$$

$$(3/2)^{665} > 2^{389} \text{ by about } 0.074 \text{ cents.}$$

Leonardo da Vinci *ca.* 1470 observed that "two men shouting together do not seem to produce twice the amount of noise that one man would" [1] and in general we now know that sensations vary as the logarithm of the stimulus (Fechner's Law). We talk and think of two octaves as twice the size of one (as on the piano keyboard), three octaves as three times the size, and so on. Yet the frequencies of these intervals are in the ratio 2:4:8 · · · . Base two logarithms are therefore naturally suited for musical purposes and were published in 1670, fifty-six years after Napier's tables [2]. Modern tables are available [3].

2. Lesser divine ratios. Over the centuries musical opinion has been remarkably consistent—

(a) The satisfying intervals are derived from natural harmonics, the frequencies of which are related as the natural number series 1:2:3 · · · .

(b) Successive ratios are favoured and are named 'superparticular.' They are an infinite series 2/1, 3/2, 4/3, 5/4, · · · .

(c) The lower members are pleasing; the higher members tend to harshness and eventually become unacceptable.

(d) Certain ratios, although within the range of acceptable harshness, are regularly rejected, e.g., 7/6, 8/7, 11/10, 12/11, 13/12, 14/13, · · · .

There is no obvious reason for this last empirical fact. However, an analysis of a large amount of material discloses that *the ear prefers superparticular ratios that are derived from the first three primes*, and when other ratios are omitted we are left with the following *finite* series of well-known intervals:

2/1	octave	9/8	major tone
3/2	perfect fifth	10/9	lesser tone
4/3	perfect fourth	16/15	diatonic semitone
5/4	major third	25/24	chromatic semitone
6/5	minor third	81/80	comma of Didymus.

The enthusiast will no doubt relate these intervals (excluding the octave) to the nine Platonic and Kepler-Poinsot regular polyhedra. Since the perfect fifth and the major third contain the first three primes, all other intervals may be compounded from them.

3. Just tunings. Perfection in tuning is an *ignis fatuus* which philosophers and musicians have followed since the beginning of time. They have concentrated on tunings largely composed of primary intervals (the 3/2 fifth and the 5/4 third) and which are loosely termed 'just tunings.' Complexity, vagueness and the absence of a simple method for recording observations have caused confusion and reduplication, but with a simple definition and a geometrical analogy suggested by T. H. O'Beirne [4] I hope to show that there are 118 just tunings and all possess undesirable qualities in varying degree.

A just tuning is one in which every note is related to at least one other note by a primary interval. Such a tuning can be plotted on squared paper. Vertices represent notes, horizontal lines joining them (left-right) perfect fifths, and vertical lines (up-down) major thirds. The problem resolves itself into finding the total number of unbroken patterns that can be formed.

Patterns are easily memorised and each completely defines a tuning. We start with the simplest—Pythagorean tuning—a sequence of eleven perfect fifths which can be plotted on a single horizontal line (Figure 3, i). The twelfth fifth uniting the last note with an octave of the first is left blank, and this indicates that it is imperfect. It is called the **Procrustean fifth** since it is cut to fit, and in this instance it is a perfect fifth less a Pythagorean comma: $2^{18}/3^{11}$. See [5].

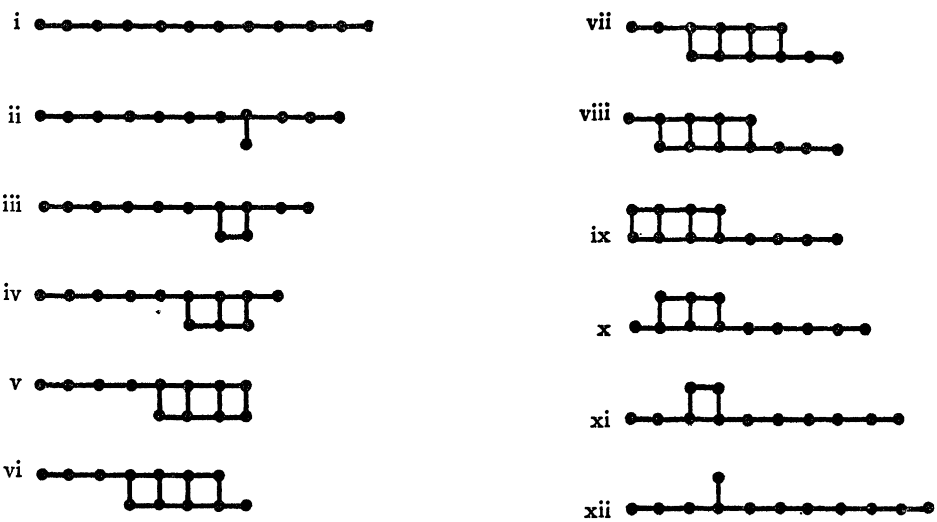


FIG. 3

Next we list all possible patterns occupying two horizontal lines (Figure 3, ii–xii). The symmetrical pattern (vii) with ten perfect fifths and four major thirds was suggested by Ramis de Pareja in 1482.

There are 43 patterns occupying three horizontal lines, and space will not allow these to be listed. The symmetrical pattern (Figure 4) is of special interest. I regard it as the most perfect of all the just tunings because it contains the maximum number of primary intervals (nine perfect fifths, eight major thirds).

Four horizontal lines give 55 patterns. The following (Figure 5) by Marpurg 1776, is generally spoken of as the 'model form' of just tuning, although it has one less perfect fifth than the symmetrical pattern above.

Five horizontal lines give 8 patterns and this completes the list of 118 just tunings.

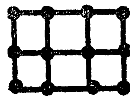


FIG. 4

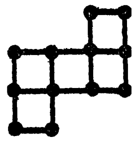


FIG. 5

Just tunings are pleasing, and each key has a character which can be suited to the mood of the composition (now a forgotten artistic refinement). Their inherent imperfections render them unacceptable for the harmonic and modulatory demands of modern music.

4. **Temperaments.** Unpleasant intervals cannot be abolished and improvement is only obtained by "tempering" or adjusting, so that the unpleasantness is shared with other intervals. This can be done in an infinite number of ways of which two will be outlined.

EQUAL TEMPERAMENT (ET). The Chinese were concerned with the problem of dividing the octave into twelve equal intervals more than 1000 years B.C. Their music did not require twelve chromatic notes but they realised the need for this number for the purpose of transposition. This meant finding the twelfth root of two which was not an easy problem. The astronomer Ho-Cheng-Tien was accused of "doing violence to figures" when he tried to find a solution *ca.* 420 A.D. Wang-Po, a physician, produced inaccurate results *ca.* 938 A.D., and not until 1598 did Prince Chu-Tsai-Yu "after meditating for days and nights before the light of Truth was revealed" come up with an answer said to be correct to nine places. In Europe the same feat was performed in 1600 by Simon Stevin, an inspector of canals in Holland, author of *La Disme,* and inventor of a sailing barge.

There is no evidence that J. S. Bach (1685–1750) attempted or intended to tune equally. The "48" were written for "*Das wohltemperierte Clavier*"—the *well tempered* clavier, not the equally tempered. It has been pointed out that the frets on ancient instruments appear to be spaced equally (in the logarithmic sense) and the 6-string lute in *The Ambassadors* by Hans Holbein the Younger 1533, has been quoted as an example. This instrument, and a number of curious objects including a German arithmetic book, lie on the lower shelf of the buffet on which the ambassadors are leaning. The finger board is foreshortened by perspective, and all in all the example is not convincing.

Equal temperament was not generally adopted until the beginning of the present century. It is a tedious temperament for the tuner because every interval is "out of tune." Accuracy is seldom achieved and then only by counting twelve different beat rates or by utilising apparatus such as the "Stroboconn."

MEANTONE TEMPERAMENT. In 1523 Aron suggested that fifths should be tempered to produce 5/4 thirds. Four perfect fifths—say C-G-D-A-E—produce a third C-E (plus two octaves) with an unpleasantly large ratio, i.e., $(3/2)^4$

divided by 4 to get rid of the octaves $=81/64$. In order to give a $5/4$ third, the ratio of each fifth must be reduced to $\sqrt[4]{5}$.

Meantone tuning is not just, because the network is broken (Figure 6). The middle note of the major third divides the latter into two equal major seconds, hence the name "meantone."

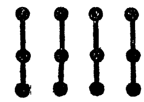


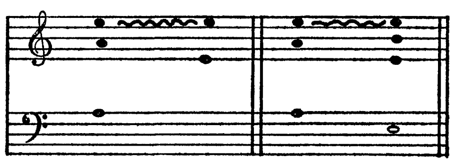
FIG. 6

This temperament was established by about 1600 and remained popular for a long time. Many organs were still tuned to it at the beginning of the present century. It died a lingering death because musicians were strongly opposed to its replacement by equal temperament.

5. Equal beating scale (EBS). This is evolved in a different manner and is not a tuning or a temperament. It possesses the following advantages:

i. It can be used for all musical purposes.
ii. It introduces a soupçon of colour to all keys.
iii. It may represent a close approximation to J. S. Bach's "well-tempered" scale.
iv. It enhances the resonance of stringed keyboard instruments.
v. Above all, ease in tuning is marked and greater accuracy is likely to be achieved.

The principle is simple. All intervals in the tuning series (Figure 2) have the same beating rate. Beats occur between the 3rd partial of the lower note and the 2nd partial of the upper note of an imperfect fifth (Figure 7, a): and between the 3rd partial of the upper note and the 4th partial of the lower note of an imperfect fourth (Figure 7, b).



(a) (b)

FIG. 7

The beat rate is the difference between the frequencies of these partials. In this example, if the frequency of the lower E is half the rate of the upper E (i.e., they are "in tune") then the beat rate is the same in each case.

The common beat rate (β) for the EBS is found by solving the twelve chain equations of the tuning series (Figure 2) in terms of β and α.

Thus

$$
\begin{array}{llll}
\text{(i)} & \beta = 3a - 2e & \text{or} & e = \tfrac{3}{2}a - \tfrac{1}{2}\beta \\
\text{(ii)} & \beta = 3e - 4b & & b = \tfrac{3}{4}e - \tfrac{1}{4}\beta \\
\text{(iii)} & \beta = 3b - 2f\# & & f\# = \tfrac{3}{2}b - \tfrac{1}{2}\beta \\
\cdots & \cdots & \cdots & \cdots \\
\text{(xii)} & & & 2a = \tfrac{3}{2}d - \tfrac{1}{2}\beta.
\end{array}
$$

From which $\beta = (7153/1568693)a$. This is the key of all necessary calculations. The ratios of all intervals can be expressed as integral numbers, and with international tuning frequency $A_4 = 440$ cps the beat rate is 1.00317 p/s which, for most practical purposes, may be taken as unity.

6. Apologia. It has been said, with regard to musical problems, that musicians generally give the correct answers supported by illogical argument, but mathematicians arrive at incorrect answers through a process of irrefutable reasoning.

This puts me in a quandary. I would like to be thought of as the operator "little i"—neither one thing nor the other, perhaps imaginary but sometimes useful.

References

1. The Notebooks of Leonardo da Vinci I, London, 1954.
2. Murray Barbour, Musical logarithms, Scripta Mathematica, 3 (1940) 21.
3. Silver and Newby, A Table of Logarithms to Base Two, TN/Math III, Ministry of Aviation, London, W.C.2, 1964.
4. T. H. O'Beirne, Puzzles and paradoxes, New Scientist, 274 (Feb. 15, 1962).
5. A. L. Leigh Silver, Notes on the Duodecimal Division of the Octave, Institute of Musical Instrument Technology, London, (Dec. 1964).
6. ———, Equal beating chromatic scale, J. Acoustical Soc. Amer., 29 (1957) 476–481.
7. ———, Some musico-mathematical curiosities, Math. Gazette, 48 (1964) 1–17.

ONTO ENDOMORPHISMS ARE ISOMORPHISMS

MORRIS ORZECH, Queen's University

1. Introduction. In this paper we shall discuss a theorem of Vasconcelos, and the following generalization thereof:

THEOREM 1. *Let M be a finitely generated R-module over the commutative ring R. Let N be any R-submodule of M. Let $f: N \to M$ be an R-module epimorphism. Then f is an isomorphism.*

Vasconcelos's result covers the case when $N = M$. In [8] it is stated more generally, for modules M whose localizations $M_{\mathfrak{m}}$ are finitely generated $R_{\mathfrak{m}}$-modules for each maximal ideal $\mathfrak{m}$ of R. But since f is onto (respectively one-one)

if and only if each $f_{\mathfrak{m}}$ is onto (respectively one-one), even if M is not finitely generated, the special case of Theorem 1 when $M=N$ implies Vasconcelos's result. Similarly, Theorem 1 itself admits a generalization to the case when M is locally finitely generated.

In Section 2 we shall derive some consequences of Theorem 1. We shall begin Section 3 by indicating several methods of approaching the proof of Vasconcelos's theorem. Two of these methods, both known by Vasconcelos, have in common the use of the theory of determinants over a commutative ring. We shall show that Theorem 1 can be proved without the use of determinants. This will be accomplished by using a well-known reduction technique, and elementary properties of noetherian rings and modules. This technique will be applied to prove a theorem analogous to the Cayley-Hamilton theorem. Other applications of this reduction technique will be mentioned in Section 4.

2. Consequences of Theorem 1. The next two results are contained in [3], and are immediate consequences of Theorem 1.

COROLLARY 1. *Let R be a commutative ring, M a free R-module of finite rank and N a free R-submodule of M. Then* $\operatorname{rank}_R N \leqq \operatorname{rank}_R M$.

COROLLARY 2. *Let R be a commutative ring and M an R-module generated by n elements. Suppose N is a free R-submodule of M with* $\operatorname{rank}_R N \leqq n$. *Then M is a free R-module and* $\operatorname{rank}_R N = \operatorname{rank}_R M$.

The result below is well known in the special case when x and y are $n \times n$ matrices over R [4, p. 158]:

COROLLARY 3. *Let R be a commutative ring and A an R-algebra which is a finitely generated R-module. If x and y are elements of A for which $xy=1$, then $yx=1$.*

3. The proof of Theorem 1. Vasconcelos has published two proofs of the case of Theorem 1 for which $M=N$. Each uses determinants. The first, in [8], involves a reduction to the case when M is free. The other proof, in [9], makes use of the following proposition, which we shall use later. The setting necessary to prove Vasconcelos's theorem is $S=R[f]$, where $f: M \to M$ is onto, and $I=Sf$.

PROPOSITION 1. *Let I be an ideal in the commutative ring S. Let M be a finitely generated S-module for which $IM=M$. Then there exists an element c in I for which $(1-c)M=0$.*

We shall comment further on this result toward the end of Section 3. At this point we shall introduce the technique of reducing to the noetherian case. The key fact needed is Hilbert's Basis Theorem, which states that if R is a noetherian ring, so is $R[X]$.

LEMMA 1. *Let S be any ring and P an S-module satisfying the ascending chain condition on submodules. Let Q be an S-submodule of P and $g: Q \to P$ an S-module map onto P. Then g is an isomorphism.*

Proof. Let $K_0=\ker(f)$ and define K_i by $K_i=g^{-1}(K_{i-1})$, $i>0$. An easy induction argument shows that $K_i\subset K_{i+1}$. The ascending chain of submodules K_i of P must terminate; thus $K_n=K_{n+1}$ for some n. If $g(x)=0$ we can choose elements x_i in K_i satisfying $x_0=x$, $g(x_{i+1})=x_i$. Since x_{n+1} is in K_n, successive applications of g yield that $x=0$.

Proof of Theorem 1. Let $f:N\to M$ be an onto map, and y_0 any nonzero element of M. We shall show that $f(y_0)\neq 0$. Let $x_1, \cdots, x_n$ be a generating set for M. Suppose $f(y_i)=x_i$ for $i=0, 1, \cdots, n$. For suitable a_{ij} and b_{ij} in R we have

$$f(x_i) = \sum_{j=1}^{n} a_{ij}x_j, \qquad y_i = \sum_{j=1}^{n} b_{ij}x_j.$$

Let F be the subring of R generated by 1; F is Z/mZ for some integer m. Let S be the subring of R generated by F and all the a_{ij} and b_{ij}:

$$S = F[a_{ij}, b_{ij}] \quad i, j = 0, 1, \cdots, n.$$

The ring S is noetherian since it is a homomorphic image of a polynomial ring over the noetherian ring F:

$$S = F[X_{ij}, Y_{ij}]/(X_{ij} - a_{ij}, Y_{ij} - b_{ij}).$$

Let $P=Sx_1+\cdots+Sx_n$ and $Q=Sy_0+\cdots+Sy_n$. Then Q is an S-submodule of P, since each generator of Q is an S-linear combination of the generators for P. The restriction of f to Q induces an S-homomorphism $g:Q\to P$; in fact g is an epimorphism since each generator of P is in $f(Q)$. Thus P is a finitely generated module over the noetherian ring S, and therefore satisfies the ascending chain condition on S-submodules. By Lemma 1, g is an isomorphism. But Q contains the element y_0, so $f(y_0)=g(y_0)\neq 0$. This completes the proof.

An alternate proof of Theorem 1 can be given by using Vasconcelos's theorem. Assume that the case of Theorem 1 when $M=N$ is true. Perform the same reduction as above, to the noetherian case. Then let $T_0=Q$, $T_{i+1}=g^{-1}(T_i)$. Let T be the intersection of all T_i. Since P is noetherian, so is the S-submodule T. But g restricts to an onto map from T to T and g is therefore one-one on T. It is clear though that $\ker(g)\subset T$.

The technique used above can be applied to prove Proposition 2, a known analogue to the Cayley-Hamilton theorem.

LEMMA 2. *Let R be a commutative noetherian ring, M and N finitely generated R-modules. Then* $\mathrm{Hom}_R(M, N)$ *is a finitely generated R-module.*

Proof. Let F be a free R-module of finite rank mapping onto M via $j:F\to M$. Then j^* is a one-one map from $\mathrm{Hom}_R(M, N)$ to $\mathrm{Hom}_R(F, N)$. The latter module is finitely generated since it is isomorphic to a direct sum of r copies of N, where $F=R^r$. But R is noetherian and thus $\mathrm{Hom}_R(M, N)$ is finitely generated, since it is a submodule of a finitely generated R-module.

PROPOSITION 2. *Let R be a commutative ring, M a finitely generated R-module and f an R-module endomorphism of M. Then f satisfies a monic polynomial with coefficients in R.*

Proof. Let $x_1, \cdots, x_n$ generate M, and let $f(x_i) = \sum_j a_{ij}x_j$. Let S be the subring of R generated by 1 and by the a_{ij}'s, and let

$$N = Sx_1 + \cdots + Sx_m.$$

S is a noetherian ring and N is thus noetherian. Let g denote the restriction of f to N; g is an S-endomorphism of N.

Now g is an element of the ring $\mathrm{End}_S(N)$, and the latter is a finitely generated S-module by Lemma 2; but

$$S[g] = \sum_{i=0}^{\infty} Sg_i$$

is a submodule of $\mathrm{End}_S(N)$ and is therefore noetherian. The chain of submodules

$$S_k = S + Sg + \cdots + Sg_k$$

must stabilize, and g^n must lie in S_{n-1} for some n. That is, for some $a_0, \cdots, a_{n-1}$ in S,

$$g^n + a_{n-1}g^{n-1} + \cdots + a_0 = 0.$$

Since this equation holds on N, it holds on $x_1, \cdots, x_m$, and thus on M with f replacing g.

COROLLARY 4. *Let R be a commutative ring and let M be a finitely generated R-module. Suppose that b in R satisfies $bM = M$. Then there is an element a in R satisfying $abx = x$ for all x in M.*

Proof. The homomorphism $f: M \to M$ sending x to bx is onto by assumption, and must be an isomorphism by Theorem 1. Let g be the inverse of f. By the proposition above we have an equation:

$$g^n + a_{n-1}g^{n-1} + \cdots + a_0 = 0.$$

Multiplying by f^n, or equivalently by b^n, yields:

$$x = -(a_{n-1} + a_{n-2}b + \cdots + a_0b^{n-1})bx$$

for x in M. This completes the proof.

We have proved Theorem 1 without use of determinants. Moreover, Proposition 1 can be proved from Theorem 1. This can be done via a reduction to the noetherian case and induction on the number of generators of I. The case for I principal is Corollary 4 above. The crucial case for $I = (a, b)$ is unpleasantly complicated, and we therefore omit the determinant-free proof of Proposition 1.

LEMMA 3. *Let R be a commutative ring, and $f: M \to N$ a homomorphism of R-*

modules, with $N/f(M)$ a finitely generated R-module. Suppose that for each maximal ideal $\mathfrak{m}$ of R, the induced map

$$f[\mathfrak{m}]: M/\mathfrak{m}M \to N/\mathfrak{m}N$$

is onto. Then f is onto.

Proof. For $\mathfrak{m}$ any maximal ideal of R we have $f(M)+\mathfrak{m}N=M$. This shows that $\mathfrak{m}(N/f(M))=N/f(M)$. By Proposition 1, there exists c in $\mathfrak{m}$ satisfying $(1-c)(N/f(M))=0$. Thus the annihilator of $N/f(M)$ is not contained in any maximal ideal $\mathfrak{m}$ of R, so $N/f(M)=0$. This completes the proof.

The following result is a trivial generalization of a theorem about central separable algebras [1, Corollary 3.4]:

PROPOSITION 3. *Let A be an R-algebra which is a finitely generated R-module. Assume that for any maximal ideal $\mathfrak{m}$ of R, $\mathfrak{m}A$ is a maximal two-sided ideal of A. Then any R-algebra endomorphism of A is an isomorphism.*

Proof. Let $f:A\to A$ be an R-algebra homomorphism. Then

$$f[\mathfrak{m}]:A/\mathfrak{m}A\to A/\mathfrak{m}A$$

is one-one, since $A/\mathfrak{m}A$ contains no proper two-sided ideals. But $A/\mathfrak{m}A$ is a finite-dimensional vector space over $R/\mathfrak{m}$ and $f[\mathfrak{m}]$ is therefore an isomorphism. By Lemma 3, f is onto and must be an isomorphism by Theorem 1. This completes the proof.

4. The reduction technique. The method used in Section 3 is well known. Strooker used it in [6, p. 750]. It is useful in proving results about finitely generated projective modules, since the noetherian ring generated by 1 has nice properties [2, pp. 476–477]. Swan uses the method to derive the result below from the same result proved when I is nilpotent rather than nil:

PROPOSITION 4. *If R is a ring and I is an ideal of R consisting of nilpotent elements, then any idempotent of R/I is the image of an idempotent in R.*

Finally, the technique is used in [5] to prove a result analogous to Vasconcelos's theorem:

PROPOSITION 5. *Let R be a commutative ring, and A a commutative R-algebra which is finitely generated as an R-algebra. Let f be an R-algebra endomorphism of A onto A. Then f is an isomorphism.*

References

1. M. Auslander and O. Goldman, The Brauer group of a commutative ring, Trans. Amer. Soc., 97 (1960) 367–409.

2. H. Bass, Algebraic K-theory, Benjamin, New York, 1968.

3. H. Flanders, Tensor and exterior powers, J. of Algebra, 7 (1967) 1–24.

4. N. H. McCoy, Rings and ideals, Carus Monograph 8, MAA, Open Court, LaSalle, Ill.

5. M. Orzech and L. Ribes, On residual finiteness and the Hopf property in rings, J. of Alg., 15 (1970) 81–88.

6. J. R. Strooker, Lifting projectives, Nagoya Math. J., 27 (1966) 747–751.
7. R. G. Swan, Algebraic K-theory, Springer Lecture Notes in Mathematics, Vol. 76, Springer-Verlag, Berlin, 1968.
8. W. V. Vasconcelos, On local and stable cancellation, An. Acad. Brasil. Ci., 37 (1965) 389–393.
9. ———, On finitely generated flat modules, Trans. Amer. Math. Soc., 138 (1969) 505–512.

CHARACTERIZATION OF KURATOWSKI 14-SETS

ERIC LANGFORD, University of Maine

The Kuratowski closure-and-complement problem asserts that in any topological space, no more than 14 sets can be formed from a given set by using only the operations of closure and complementation, iterated in any order; moreover, the problem asserts that there exist examples of sets where the full range of 14 is actually attained. (See Kuratowski [5].) We shall call such a set a *14-set*. A standard example of a 14-set on the real line is $(0, 1)\cup(1, 2)\cup Q(2, 3) \cup\{4\}$, where $Q(2, 3)$ denotes the set of rationals in $(2, 3)$. This paper gives necessary and sufficient conditions that a subset X of an arbitrary topological space S be a 14-set; the conditions become simpler if the space S is connected. This allows us to answer fully a problem of S. Baron, concerning 14-sets on the real line [2].

Consider an arbitrary topological space S and any subset X of S. Let k, c, b, and i denote the complementation, closure, boundary, and interior operators respectively; we shall write them on the right. We shall make much use of the well-known facts that $Xi = Xkck$ and $Xb = Xkb = Xc\cap Xkc$.

The Kuratowski problem is solved as follows: Suppose that G is open, i.e., $G = Gi$. Then $G\subseteq Gc$ so that $G = Gi\subseteq Gci$ and hence $Gc\subseteq Gcic$. But $Gci\subseteq Gc$ and hence $Gcic\subseteq Gcc = Gc$. Thus $Gc = Gcic = Gckckc$. But $Xi = Xkck$ is always open, so that $Xkckc = Xkckckckc$. Similarly, $Xki = Xck$ is open, so that $Xckc = Xckckckc$. Thus the 14 possible sets are given by the following:

$$Xckckck,\ Xckckc,\ Xckck,\ Xckc,\ Xck,\ Xc,\ X,\ Xk,\ Xkc,$$
$$Xkck,\ Xkckc,\ Xkckck,\ Xkckckc, \text{ and } Xkckckck.$$

Let us denote this family of sets by $CK(X)$. Clearly, X is a 14-set iff these are all distinct; further, it is obvious that $CK(X) = CK(Xk)$ so that X is a 14-set iff Xk is a 14-set.

It is not hard to see that the 14 sets in $CK(X)$ fall naturally into two groups of seven sets each. The first group consists of $Xcic$, Xci, Xc, X, Xi, Xic, and $Xici$, and the second group consists of $Xcick = Xkici$, $Xcik = Xkic$, $Xck = Xki$, Xk, $Xik = Xkc$, $Xick = Xkci$, and $Xicik = Xkcic$. As was shown by Chapman [3], it follows that a maximum of seven sets can be formed from a given set X by using only the operations of closure and interior, namely those sets in the first group. We shall denote the family of these seven sets by $CI(X)$. From the

list above, we see that $CI(Xk)$ consists precisely of the complements of the sets in $CI(X)$; moreover, $CK(X) = CI(X) \cup CI(Xk)$.

A good deal of work has been done on the interrelationships of the sets in $CI(X)$. N. Levine [6] gives necessary and sufficient conditions on X in order that $Xci = Xic$; his work is generalized by Chapman [3], who investigates the $\binom{7}{2} = 21$ possible equalities between sets in $CI(X)$ and gives necessary and sufficient conditions that any given equality hold. Aull [1] calls a set X an *n-set* ($n = 1, 2, \cdots, 7$) if $CI(X)$ contains precisely n distinct sets. He calls a topological space S an *n-space* if S contains an n-subset but no $(n+1)$-subset. He then characterizes n-spaces; the most thorough analysis is for Hausdorff spaces. Herda and Metzler [4] investigate the same problem for finite spaces and show that every 7-space must have at least seven points.

The visualization of the containments that must occur between the various sets in $CI(X)$ is most easily made by considering the following diagram (Figure 1) due to Kuratowski [5]. He showed that if a set $X\tau$ is accessible from a set $X\sigma$ by following arrows, then $X\sigma \subseteq X\tau$. Conversely, he showed that if no such path exists, then there exist examples where containment does not hold. This diagram will be most helpful later on. A similar diagram holds for the sets in $CI(Xk)$.

$$\begin{tikzcd}
& & Xci \arrow[dr] & & \\
Xi \arrow[r] \arrow[ddrr] & Xici \arrow[ur] \arrow[dr] & & Xcic \arrow[r] & Xc \\
& & Xic \arrow[ur] & & \\
& & X \arrow[uurr] & &
\end{tikzcd}$$

FIG. 1

Aull [1] remarks without proof that X is a 14-set iff $CI(X)$ contains precisely seven sets, i.e., iff X is a 7-set in his terminology. Our proof of Theorem 1 makes use of this fact; since no proof seems to occur in the literature, we shall include one here for completeness.

LEMMA 1. *A subset X of the topological space S is a* 14-*set iff precisely seven distinct sets can be formed from X by using the operations of closure and interior, iterated in any order.*

Proof. Certainly X is a 7-set whenever it is a 14-set. To show the converse, assume to the contrary that X is a 7-set but not a 14-set. Now $CI(X)$ consists of precisely seven sets; since $CI(Xk)$ consists of the complements of the sets in $CI(X)$, evidently $CI(Xk)$ must consist of seven distinct sets also. Since X is not a 14-set, some set in $CI(X)$ must be the same as some set in $CI(Xk)$; that is, two of the sets in $CI(X)$ must be complements of each other. From Figure 1, it is seen that Xc contains as a subset every set in $CI(X)$; since Xc contains both a set and its complement, it follows that $Xc = S$. But then $Xcic = Sic = S = Xc$, and X cannot be a 7-set. This contradiction shows that X is a 14-set whenever it is a 7-set.

The following theorem is the main theorem of the paper.

THEOREM 1. *Let X be a subset of the arbitrary topological space S. Then X is a 14-set iff the following five conditions hold:*

A. $Xbi = Xkbi = Xci \setminus Xic \neq \varnothing$
B. $X \cap Xckckck = X \setminus Xcic \neq \varnothing$
C. $Xk \cap Xkckckck = Xici \setminus X \neq \varnothing$
D. $Xcib = Xcic \setminus Xci \neq \varnothing$
E. $Xkcib = Xic \setminus Xici \neq \varnothing$.

Moreover, the conditions are independent, i.e., in general, no four imply the fifth.

Proof. Note that B and C are dual, that D and E are dual, and that A is self-dual. We show first the necessity of the five conditions by showing that if any fails, then X is not a 14-set.

Suppose first that A fails, so that $Xci \subseteq Xic$. Then it is easily shown that $Xci = Xici$ and $Xic = Xcic$, so that X is not a 14-set. (This is proved as Lemma 1 of [4].) Suppose then that B fails, so that $X \subseteq Xcic$. From Figure 1, we see that $X \subseteq Xcic \subseteq Xc$; since $Xcic$ is closed and pinned in between X and its closure, we must have that $Xcic = Xc$ and X cannot be a 14-set. Similarly if C fails, then $Xi = Xici$ and X is not a 14-set. Finally, if D fails, then $Xcic = Xci$ and if E fails, then $Xic = Xici$, so that each of the conditions is necessary.

Now we shall show that the five conditions are sufficient. Select five points as follows:

$$a \in Xci \setminus Xic = Xbi$$

$$b \in X \setminus Xcic = X \cap Xckckck$$

$$c \in Xici \setminus X = Xk \cap Xkckckck$$

$$d \in Xcic \setminus Xci = Xcib$$

$$e \in Xic \setminus Xici = Xkcib.$$

Now consider the following table, where a Y entry means that the point lies in the set and an N entry means that it does not. A blank space indicates that the point may or may not lie in the set. The table can be filled in by examination of the inclusions given in Figure 1.

It can now be verified by direct inspection that the seven sets in $CI(X)$ are all distinct, since given any pair of them, at least one of the five points lies in one of the sets, but not the other. Since these sets are distinct, it follows from Lemma 1 that X is a 14-set.

To show that no four of the conditions imply the fifth, we shall give examples; we need not go outside of the real line.

Example 1. All properties except A.

Let $S = R$ and $X = \{0\} \cup (2, 3) \cup (3, \infty)$. Then $Xb = \{0, 2, 3\}$ and $Xbi = \varnothing$, yet $X \cap Xckckck = \{0\}$, $Xk \cap Xkckckck = \{3\}$, and $Xcib = Xkcib = \{2\}$.

	a	b	c	d	e
$Xcic$	Y	N	Y	Y	Y
Xci	Y	N	Y	N	
Xc	Y	Y	Y	Y	Y
X		Y	N		
Xi	N	N	N	N	N
Xic	N	N	Y		Y
$Xici$	N	N	Y	N	N

TABLE I

Example 2. All properties except B.

Let $S=R$ and let $X=Q^{-}\cup(1,\ 2)\cup(2,\ \infty)$, where Q^{-} is the set of negative rationals. Then $Xbi=(-\infty,\ 0)$, $Xckckck=(0,\ 1)$, so that $X\cap Xckckck=\varnothing$, yet $Xk\cap Xkckckck=\{2\}$, $Xcib=\{0,\ 1\}$, and $Xkcib=\{1\}$.

Example 3. All properties except C.

Let $S=R$ and let X be the complement of the set in Example 2.

Example 4. All properties except D.

As we shall show in the next theorem, conditions A, B, and C imply conditions D and E when S is connected. For our example, we shall disconnect R by removing a single point. Let $S=R\setminus\{2\}$ under the usual relative topology, and let $X=Q^{-}\cup(0,\ 1)\cup(1,\ 2)\cup\{3\}$. Then $Xbi=(-\infty,\ 0)$, $X\cap Xckckck=\{3\}$, $Xk\cap Xkckckck=\{1\}$, and $Xkci=(-\infty,\ 0)\cup(2,\ \infty)$, so that $Xkcib=\{0\}$. But $Xci=(-\infty,\ 2)$ and hence $Xcib=\varnothing$. (Note that Xci is both open and closed in S.)

Example 5. All properties except E.

Let $S=R\setminus\{2\}$ and let X be the complement of the set in Example 4.

In Example 4 in the proof of Theorem 1, we remarked that if S were connected, then D and E were consequences of A, B, and C. We prove this now.

THEOREM 2. *If S is a connected space, then a subset X of S is a* 14-*set iff* A, B, *and* C *hold.*

Proof. The conditions are certainly necessary. We show they are sufficient by showing that since S is connected, A, B, and C imply both D and E.

Suppose that A, B, and C are true, but that D is false, i.e., that $Xcib=\varnothing$ so that $Xci=Xcic$. This means that Xci is both open and closed in S. Since S is connected, this implies that either $Xci=\varnothing$ or $Xci=S$. If $Xci=\varnothing$, then A is violated, whereas if $Xci=S$, then $Xcic=Sc=S$, so that B is contradicted. Hence A and B imply D. In a similar fashion, we can show that A and C imply E. Thus A, B, and C imply D and E, so that X is a 14-set by Theorem 1.

Let us examine conditions B and C more closely. We note that in each of the examples in Theorem 1 it is true that $X\cap Xckckck$ consists of precisely the isolated points of X. This is not true in general; however, we do have the following interesting characterization of the set $X\cap Xckckck$.

THEOREM 3. *Let S be an arbitrary topological space and suppose that X is a subset of S. Then the set $X\cap Xckckck$ is the largest relatively open subset of X which is nowhere dense in S.*

Proof. Recall that a set Y is *nowhere dense* if $Yci = Yckck = \emptyset$. Certainly $X\cap Xckckck$ is a relatively open subset of X. Consider then $(X\cap Xckckck)ci$; we have that $(X\cap Xckckck)c\subseteq Xc\cap Xckckckc = Xc\cap Xckc$, so that

$$(X \cap Xckckck)ci \subseteq (Xc \cap Xckc)i = Xci \cap Xckci = Xckck \cap Xckckck.$$

But $Xckck\subseteq Xckckc$ so that $Xckck\cap Xckckck = \emptyset$. Thus $X\cap Xckckck$ is a relatively open subset of X which is nowhere dense in S.

Let us suppose now that Y is any relatively open subset of X which is nowhere dense in S. We shall show that $Y\subseteq X\cap Xckckck$. If $Y=\emptyset$, there is nothing to prove, so assume that $Y\neq\emptyset$. Write $Y=X\cap G$, where G is open in S; certainly $G\neq\emptyset$. Consider the set $H=G\cap Yck$. Now H is open, and if $H=\emptyset$, then $G\subseteq Yc$, which is impossible since G is nonempty and Y is nowhere dense; thus $H\neq\emptyset$. Since $Y=G\cap X\subseteq Yc$, certainly $\emptyset = Y\cap Yck = G\cap X\cap Yck = H\cap X$, so that $H\subseteq Xk$. But H is open, so that $H\subseteq Xki = Xck$, and thus $H\subseteq Xckc$.

We now show that $Y\subseteq Xckc$. Suppose that $y\in Y$ and assume that N is any neighborhood of y; we shall show that N meets Xck. Let $M=N\cap G$, which is nonempty since $y\in Y\subseteq G$. Moreover, M is open and $M\cap Yck\neq\emptyset$ since Y is nowhere dense. Thus $\emptyset\subset M\cap Yck = N\cap G\cap Yck = N\cap H\subseteq N\cap Xck$, so that $N\cap Xck\neq\emptyset$, and therefore $y\in Xckc$. Since y was arbitrary, it follows that $Y\subseteq Xckc$ and hence $Yc\subseteq Xckc$. Thus it follows that $Xckc\supseteq H\cup Yc = (G\cap Yck)\cup Yc = G\cup Yc\supseteq G$. Since G is open, we have that $G\subseteq Xckci = Xckckck$. But then $Y=G\cap X\subseteq X\cap Xckckck$, as was to be shown.

In light of the last theorem, we can rephrase conditions B and C in Theorems 1 and 2 as follows:

B*. X contains a nonempty relatively open subset which is nowhere dense in S.

C*. Xk contains a nonempty relatively open subset which is nowhere dense in S.

Recall that $x\in X$ is an *isolated point* of X if there exists a neighborhood N of x such that $N\cap X = \{x\}$. The following lemma is immediate:

LEMMA 2. *Suppose that S is a space in which points (i.e., singletons) are closed but not open. (For example, we could assume that S is T_1 and connected.) Then every singleton set is nowhere dense in S. In particular, if X is a subset of S and if x is an isolated point of X, then $\{x\}$ is a relatively open subset of X which is nowhere dense in S so that $x\in X\cap Xckckck$ by Theorem 3.*

Thus in any T_1 connected space, and in particular, on the real line, the existence of isolated points in X will guarantee that $X \cap Xckckck \neq \emptyset$. However, the converse is false, even on the real line.

Example 6. Let S be the real line and let

$$X = Q^- \cup C \cup (2, 3) \cup (3, \infty),$$

where C is the standard Cantor set on $[0, 1]$. Then $Xbi = (-\infty, 0)$, $X \cap Xckckck = C$, and $Xk \cap Xkckckck = \{3\}$, so that X is a 14-set; yet X contains no isolated points.

If we interpret condition A on the real line, we see that it is equivalent to the property that $Xb = Xc \cap Xkc$ must contain a nonempty open interval I; that is, there must exist a nonempty open interval I such that X and Xk are both dense in I. Thus we can answer Baron's problem completely as follows:

THEOREM 4. *Suppose that X is a subset of the real line under its usual topology. Then X is a 14-set iff the following: there exists a nonempty open interval I such that X and Xk are both dense in I, and both X and Xk contain nonempty, relatively open, nowhere dense subsets of R. The following condition is sufficient, but not necessary that X be a 14-set: there exists a nonempty open interval I such that X and Xk are both dense in I, and both X and Xk contain isolated points.*

References

1. C. E. Aull, Classification of topological spaces, Bulletin de l'Académie Polonaise des Sciences Math. Astron. Phys., 15 (1967), 773–778.

2. S. Baron, Problem 5569, this MONTHLY, 75 (1968) 199.

3. T. A. Chapman, A further note on closure and interior operators, this MONTHLY, 69 (1962) 524–529.

4. H.-H. Herda and R. C. Metzler, Closure and interior in finite topological spaces, Colloquium Math., 15 (1966) 211–216.

5. C. Kuratowski, L'opération Ā de l'analysis situs, Fund. Math., 3 (1922) 182–199.

6. N. Levine, On the commutativity of the closure and interior operators in topological spaces, this MONTHLY, 68 (1961) 474–477.

MATHEMATICAL NOTES

EDITED BY DAVID DRASIN

After May 1, 1971, manuscripts for this Department should be sent to Robert Gilmer, Department of Mathematics, Florida State University, Tallahassee, FL 32306.

ASYMPTOTIC DISTRIBUTION OF REAL NUMBERS MODULO ONE

J. L. BROWN, JR., AND R. L. DUNCAN, Pennsylvania State University

For a sequence $\{x_j\}_1^\infty$ of real numbers with corresponding fractional parts $\{\beta_j\}_1^\infty$, we define functions $F_n (n \geqq 1)$ on $[0, 1]$ so that $F_n(x)$ is one nth of the number of terms $\beta_1, \beta_2, \cdots, \beta_n$ in the interval $[0, x)$. Then each F_n is nonde-

creasing on $[0, 1]$, left-continuous on $(0, 1]$ and satisfies $F_n(0)=0$, $F_n(1)=1$. Given a continuous nondecreasing function F on $[0, 1]$ with $F(0)=0$ and $F(1)=1$, we say $\{x_j\}_1^\infty$ is **distributed according to** F **modulo** 1 if

$$\lim_{n\to\infty} F_n(x) = F(x) \qquad \text{for all } x \in [0, 1]. \tag{1}$$

Perhaps the best-known necessary and sufficient condition for $\{x_j\}$ to be distributed according to F modulo 1 is

$$\lim_{n\to\infty} \frac{1}{n} \sum_{j=1}^{n} e^{2\pi i\nu x_j} = \int_0^1 e^{2\pi i\nu x} dF(x) \tag{2}$$

for all $\nu \geqq 1$. See [1–4]. The purpose of the present paper is to give a simple derivation of (2) which at the same time leads to alternative criteria appearing in the literature.

We denote by C_p the class of all real-valued continuous functions f defined on $[0, 1]$ satisfying $f(0)=f(1)$. Thus any $f\in C_p$ may be extended to $(-\infty, \infty)$ as a continuous function with period one. Our first theorem gives a necessary and sufficient condition for a sequence $\{x_j\}$ to be distributed according to F mod 1, similar to a result given by G. Helmberg [5] in a more abstract setting.

THEOREM 1. *The sequence $\{x_j\}_1^\infty$ is distributed according to F mod 1 if and only if*

$$\lim_{n\to\infty} \frac{1}{n} \sum_{j=1}^{n} f(\beta_j) = \int_0^1 f(x) dF(x) \qquad \textit{for each } f \in C_p. \tag{3}$$

Proof. If $\{x_j\}_1^\infty$ is distributed according to F mod 1, then (1) holds by definition and (3) follows directly from Helly's Second Theorem [6] since

$$(1/n) \sum_{j=1}^{n} f(\beta_j) = \int_0^1 f(x) dF_n(x).$$

Conversely, assume (3) holds for every $f\in C_p$. Let $\xi\in(0, 1)$ and define the step-function V_ξ as follows:

$$V_\xi(x) = \begin{cases} 1 \text{ for } 0 \leqq x < \xi \\ 0 \text{ for } \xi \leqq x \leqq 1. \end{cases}$$

Then, given $\epsilon>0$, there exist functions $f_1, f_2\in C_p$ such that $f_1(x)\leqq V_\xi(x)\leqq f_2(x)$ for $x\in[0, 1]$ and

$$\int_0^1 [f_2(x) - f_1(x)] dF(x) \leqq \epsilon/2.$$

For the same ϵ, there exists an $N>0$ such that $n>N$ implies

$$\begin{aligned}(4)\qquad \int_0^1 f_1 dF - \frac{\epsilon}{2} \leqq \frac{1}{n}\sum_1^n f_1(\beta_j) &\leqq \frac{1}{n}\sum_1^n V_\xi(\beta_j)\\ &\leqq \frac{1}{n}\sum_1^n f_2(\beta_j) \leqq \int_0^1 f_2 dF + \frac{\epsilon}{2}\cdot\end{aligned}$$

But, we also have

$$(5)\qquad \int_0^1 f_2\,dF - \frac{\epsilon}{2} \leqq \int_0^1 f_1\,dF \leqq \int_0^1 V_\xi\,dF \leqq \int_0^1 f_2\,dF \leqq \int_0^1 f_1\,dF + \frac{\epsilon}{2}\cdot$$

Combining (4) and (5), we find that for $n>N$,

$$\int_0^1 V_\xi(x)dF(x) - \epsilon \leqq \frac{1}{n}\sum_1^n V_\xi(\beta_j) \leqq \int_0^1 V_\xi(x)dF(x) + \epsilon,$$

or

$$\lim_{n\to\infty}\frac{1}{n}\sum_1^n V_\xi(\beta_j) = \int_0^1 V_\xi(x)dF(x).$$

Since $(1/n)\sum_1^n V_\xi(\beta_j) = F_n(\xi)$ and $\int_0^1 V_\xi(x)dF(x) = F(\xi) - F(0) = F(\xi)$, the theorem is immediate.

A sequence of (complex-valued) continuous functions $\{\phi_\nu\}_0^\infty$ will be called **fundamental** [7] in C_p if the set of finite linear combinations of terms from the sequence is dense in C_p with respect to the uniform (Chebyshev) norm; that is, for each $f\in C_p$ and each $\epsilon>0$, there exists a linear combination $\sum_0^n \lambda_i\phi_i$ such that

$$\left\|f - \sum_0^n \lambda_i\,\phi_i\right\|_\infty \equiv \sup_{x\in[0,1]}\left|f(x) - \sum_0^n \lambda_i\,\phi_i\right| < \epsilon.$$

Our next theorem shows that we may replace the set of *all* continuous functions in C_p by a fundamental set in Theorem 1 and retain the same conclusion.

THEOREM 2. *Let $\{\phi_\nu\}_0^\infty$ be a sequence of continuous functions fundamental in C_p. Then $\{x_j\}_1^\infty$ is distributed according to F modulo 1 if and only if*

$$(6)\qquad \lim_{n\to\infty}\frac{1}{n}\sum_{j=1}^n \phi_\nu(\beta_j) = \int_0^1 \phi_\nu(x)dF(x) \qquad \textit{for all } \nu \geqq 0.$$

Proof. Necessity again follows from Helly's theorem. Conversely, assume (6) holds, and let $f\in C_p$. By hypothesis, given $\epsilon>0$, there exists an $N>0$ and coefficients $a_1, \cdots, a_N$ such that $\|f-\sum_1^N a_\nu\phi_\nu\|_\infty<\epsilon/3$. For this ϵ, this N, and $a=\max(|a_1|, \cdots, |a_N|)$, there exists from (6) an $N_1>0$ such that $n>N_1$ implies

$$\max_{1\leqq\nu\leqq N}\left|\frac{1}{n}\sum_{j=1}^n \phi_\nu(\beta_j) - \int_0^1 \phi_\nu(x)dF(x)\right| < \frac{\epsilon}{3Na}\cdot$$

Then, for $n > N_1$,

$$\begin{aligned}\left|\frac{1}{n}\sum_{j=1}^{n} f(\beta_j) - \int_0^1 f(x)dF(x)\right| &\leqq \left|\frac{1}{n}\sum_{j=1}^{n}\left[f(\beta_j) - \sum_{\nu=1}^{N} a_\nu\phi_\nu(\beta_j)\right]\right| \\ &\quad + \left|\sum_{\nu=1}^{N} a_\nu\left[\frac{1}{n}\sum_{j=1}^{n}\phi_\nu(\beta_j) - \int_0^1 \phi_\nu(x)dF(x)\right]\right| \\ &\quad + \left|\int_0^1\left[\sum_{\nu=1}^{N} a_\nu\phi_\nu(x) - f(x)\right]dF(x)\right| \\ &< \frac{\epsilon}{3} + Na\left(\frac{\epsilon}{3Na}\right) + \frac{\epsilon}{3} = \epsilon.\end{aligned}$$

Thus $\lim_{n\to\infty}(1/n)\sum_1^n f(\beta_j) = \int_0^1 f(x)dF(x)$ for each f in C_p, and $\lim_{n\to\infty} F_n(x) = F(x)$ from Theorem 1.

Application. Condition (2) is an immediate consequence of Theorem 2 since $\{e^{2\pi i\nu x}\}_{-\infty}^{\infty}$ is fundamental in C_p by Fejer's classical result ([7], p. 123). Moreover, the condition for $\nu = 0$ is satisfied identically for any $F(x)$, while satisfaction of (2) for $\nu \geqq 1$ clearly implies that (2) holds for all $\nu \leqq -1$.

The following lemma shows how fundamental sequences may be generated from L_2-sequences $\{\psi_\nu\}_0^\infty$ which are complete with respect to $L_2[0, 1]$ in the L_2 norm; that is, for $f \in L_2[0, 1]$ and $\epsilon > 0$, there exists a finite linear combination $\sum_0^n \alpha_\nu\psi_\nu$ such that $\|f - \sum_0^n \alpha_\nu\psi_\nu\|_2 < \epsilon$, where $\|f\|_2 \equiv (\int_0^1 |f(t)|^2 dt)^{1/2}$ is the usual L_2-norm. For brevity, we shall refer to such a sequence $\{\psi_\nu\}_0^\infty$ as being **complete** in L_2 $[0, 1]$.

LEMMA. *Let $\{\psi_\nu\}_0^\infty$ be a complete sequence in $L_2[0, 1]$. Then the sequence $\{1, \phi_0, \phi_1, \phi_2, \cdots\}$ is fundamental in C_p where $\phi_\nu(x) \equiv \int_0^x \psi_\nu(t)\,dt$ for $\nu \geqq 0$ and $x \in [0, 1]$.*

Proof. Let C^1 denote the class of continuously differentiable functions on $[0, 1]$ which vanish at the origin. Let $f \in C^1$; then $f' \in L_2[0, 1]$, and for $\epsilon > 0$, there exists a linear combination $\sum_0^N \beta_\nu\psi_\nu$ such that $\|f' - \sum_0^N \beta_\nu\psi_\nu\|_2 < \epsilon$. Then, using the Cauchy-Schwarz inequality,

$$\begin{aligned}\left|f(x) - \sum_0^N \beta_\nu\,\phi_\nu(x)\right| &= \left|\int_0^x\left[f'(t) - \sum_0^N \beta_\nu\,\psi_\nu(t)\right]dt\right. \\ &\leqq \left[\int_0^x\left|f'(t) - \sum_0^N \beta_\nu\,\psi_\nu(t)\right|^2 dt\right]^{1/2} x^{1/2} \\ &\leqq \left\|f' - \sum_0^N \beta_\nu\,\psi_\nu\right\|_2 < \epsilon.\end{aligned}$$

Hence $\|f - \sum_0^N \beta_\nu\,\phi_\nu\|_\infty < \epsilon$. Since all powers x^k with $k \geqq 1$ are in C^1, it is clear that any polynomial can be approximated in the uniform norm by finite linear com-

binations of terms from the sequence $\{1, \phi_0, \phi_1, \phi_2, \cdots\}$. From the Weierstrass theorem, it then follows that the designated sequence is fundamental in C_p.

THEOREM 3. *Let* $\{\psi_\nu\}_0^\infty$ *be a complete sequence in* $L_2[0, 1]$ *and define* $\phi_\nu(x) = \int_0^x \psi_\nu(t)\, dt$ *for* $\nu \geqq 0$ *and* $x \in [0, 1]$. *Then* $\{x_j\}_1^\infty$ *is distributed according to* F *modulo* 1 *if and only if*

$$\lim_{n\to\infty} \frac{1}{n} \sum_{j=1}^{n} \phi_\nu(\beta_j) = \int_0^1 \phi_\nu(x)dF(x) \qquad \text{for } \nu \geqq 0. \tag{7}$$

Proof. From the lemma, $\{1, \phi_0, \phi_1, \cdots\}$ is fundamental in C_p; the result follows from Theorem 2 since condition (6) for the unity function is vacuously satisfied and may be omitted.

COROLLARY. *Let* $\{\psi_\nu\}_0^\infty$ *and* $\{\phi_\nu\}_0^\infty$ *be as in Theorem* 3. *The statement of Theorem* 3 *is valid if for* $\nu \geqq 0$, *we replace* ϕ_ν *by* θ_ν, *where*

$$\theta_\nu(x) \equiv \int_x^1 \psi_\nu(t)dt \qquad \text{for } \nu \geqq 0 \quad \text{and} \quad x \in [0, 1].$$

Proof. Define $\gamma_\nu \equiv \int_0^1 \psi_\nu(t)\, dt$ for $\nu \geqq 0$. Then $\phi_\nu(x) = \gamma_\nu - \theta_\nu(x)$ for $\nu \geqq 0$, so that the sequence $\{1, \theta_0, \theta_1, \cdots\}$ is clearly fundamental in C_p and the result follows as in Theorem 3.

This latter result, using $\{\theta_\nu(x)\}$, is essentially the generalized Weyl criterion given in [2].

In conclusion, we note that continuity of the limit function F is not a necessary restriction. We may require F merely to be nondecreasing on $[0, 1]$ with $F(0) = 0$ and $F(1) = 1$ and then modify the definition of mod 1 distribution so that (1) is to hold only at continuity points of F (the usual weak convergence concept for distribution functions). Then Theorem 1 remains valid, but V_ξ in the proof will be defined only for those values ξ at which F is continuous (a dense set in the unit interval), so that the Stieltjes integral $\int_0^1 V_\xi(x)dF(x)$ will have meaning. The remaining results in the paper follow without essential modification.

Acknowledgement. The first author's contribution was sponsored by the Office of Naval Research, NR 041-373. Reproduction in whole or in part is permitted for any purpose of the United States Government.

References

1. I. Schoenberg, Über die asymptotische Verteilung reeller Zahlen mod 1, Math. Z., 28 (1928) 171–199.

2. J. L. Brown, Jr. and R. L. Duncan, A generalized Weyl criterion, Duke Math J., 35 (1968) 699–705.

3. N. Artemiadis and L. Kuipers, On the Weyl criterion in the theory of uniform distribution modulo 1, this MONTHLY, 76 (1969) 654–656.

4. I. Niven, Irrational Numbers, MAA Carus Mathematical Monographs, No. 11, Wiley, New York, 1956, pp. 76–81.

5. G. Helmberg, Gleichverteilte Folgen in lokal kompakten Räumen, Math. Z, 86 (1964) 157–189.

6. D. V. Widder, The Laplace Transform, Princeton University Press, Princeton, N. J., 1946, p. 31.

7. E. W. Cheney, Introduction to Approximation Theory, McGraw-Hill, New York, 1966, p. 193.

ON UNIFORM CONNECTION PROPERTIES

P. J. Collins, St. Edmund Hall, Oxford

1. Introduction. It is the purpose of this note to introduce the concepts of uniform local connection and property S, well known in the theory of metric spaces, into the theory of uniform spaces. (For example, see Hall and Spencer [**2**], pages 135 and 211.) Our definitions not only simplify and generalize those customarily given for metric spaces but, taken together with the existing theory of uniform spaces, allow simpler proofs of more general theorems. We support this assertion by using the definitions in a proof of the following theorem (cf. Whyburn [**3**], pages 20–23).

THEOREM. *In order that a compact Hausdorff space be locally connected it is necessary and sufficient that it have property S.*

Further, the definitions allow a straightforward proof (not given here) of the equivalence of property S and a suitable extension of partitionability for arbitrary Hausdorff uniform spaces. (For example, see Hall and Spencer [**2**], chapter 6.)

Section 2 below gives the new definitions and states the easily proved propositions which show that they indeed extend those for metric spaces. Section 3 states and proves three lemmas which together imply our theorem.

Throughout the note, our terminology and notation will be that of Bourbaki [**1**]. Here $\mathfrak{U}$ will denote the family of entourages of a uniform space X. A metric space will always have its usual uniformity.

2. The definitions.

DEFINITION 1. X is said to be *uniformly locally connected* if for each $U\in\mathfrak{U}$, there exists $V\in\mathfrak{U}$ such that $V\subset U$ and $V(x)$ is connected for each $x\in X$.

PROPOSITION 1. *Suppose that M is a metric space. Then M is uniformly locally connected if and only if for each $\epsilon>0$ there exists $\delta>0$ such that any two points x and y of M whose distance apart is less than δ lie together in a connected subset of M of diameter less than ϵ.*

DEFINITION 2. X is said to have *property S* if for each $U\in\mathfrak{U}$, the set X is covered by a finite family of connected U-small sets.

PROPOSITION 2. *Suppose that M is a metric space. Then M has property S if and only if for each $\epsilon>0$, the space M is the union of a finite number of connected sets, each of diameter less than ϵ.*

3. Proof of the theorem.

LEMMA 1. *A locally connected compact Hausdorff space X is uniformly locally connected.*

Proof. Let $\mathfrak{U}$ denote the unique uniformity compatible with the topology of X. Suppose that $U\in\mathfrak{U}$ and that V is a symmetric entourage such that $V\circ V\subset U$. Since X is locally connected, for each $x\in X$ there exists $C_x\in\mathfrak{U}$ such that $C_x(x)$ is connected and $C_x(x)\subset V(x)$. Since X is compact Hausdorff, the fact that

$$W=\bigcup_{x\in X}[C_x(x)\times C_x(x)]$$

is a neighbourhood of the diagonal implies that W is an entourage. As a union of connected sets containing x, each $W(x)$ is connected for $x\in X$, and we clearly have

$$W\subset\bigcup_{x\in X}[V(x)\times V(x)]\subset V\circ V\subset U.$$

LEMMA 2. *Each precompact uniformly locally connected space X has property S.*

Proof. Suppose that $U\in\mathfrak{U}$, and let V be a symmetric entourage such that $V\circ V\subset U$. Since X is uniformly locally connected, there exists $W\in\mathfrak{U}$ such that $W\subset V$ and each $W(x)$ is connected. Since X is precompact, there exists a finite family $\{A_i\}_{i=1}^n$ of nonempty W-small sets which cover X. For $1\leqq i\leqq n$, choose $x_i\in A_i$. Then clearly $\{W(x_i)\}_{i=1}^n$ is a finite family of connected sets covering X, and for $1\leqq i\leqq n$,

$$W(x_i)\times W(x_i)\subset V(x_i)\times V(x_i)\subset V\circ V\subset U.$$

LEMMA 3. *If X has property S, then X is locally connected.*

Proof. Suppose that $x\in X$, that $U\in\mathfrak{U}$, and that V is a closed entourage such that $V\subset U$. Since X has property S, there exists a finite family $\{C_i\}_{i=1}^n$ of connected V-small sets covering X. Let C be the union of those C_i such that $x\in\overline{C}_i$; thus C is connected. Then C is a neighborhood of x since x neither belongs to nor is a limit point of $X-C$; also, since V is closed,

$$C\subset\bigcup_{x\in\overline{C}_i}C_i\subset\bigcup_{x\in\overline{C}_i}\overline{C}_i\subset V(x)\subset U(x).$$

The theorem of section 1 now follows immediately from Lemmas 1, 2, and 3.

Added in proof: Our theorem tells us that for compact Hausdorff spaces our use of the term "property S" agrees with that of Wallace [4 page 98]. Other uses of the term are considered in [5, pages 580–582].

References

1. N. Bourbaki, General Topology, Part 1, Hermann, Paris, and Addison-Wesley, Reading, Mass., 1966, chapter II.

2. D. W. Hall and G. L. Spencer, Elementary Topology, Wiley, New York, 1955.
3. G. T. Whyburn, Analytic Topology, Amer. Math. Soc., 1963, Revised edition.
4. A. D. Wallace, Extensional invariance, Trans. Amer. Math. Soc., 70 (1951) 97–102.
5. M. Henriksen and J. R. Isbell, Local connectedness in the Stone-Čech compactification, Illinois J. Math., 1 (1957) 574–582.

A SIMPLE ESTIMATE FOR THE NUMBER OF STEPS IN THE EUCLIDEAN ALGORITHM

J. D. DIXON, Carleton University, Ottawa

The usual Euclidean algorithm for computing the greatest common divisor of two integers may be described as follows.

For any positive integers u, v with $u \leqq v$, we compute positive integers q_i and integers r_i such that

$$(1) \qquad r_0 = v, \quad r_1 = u \quad \text{and} \quad r_{i-1} = q_i r_i + r_{i+1} \qquad (i = 1, 2, \cdots, n)$$

with $r_0 \geqq r_1 > \cdots > r_n > r_{n+1} = 0$. We shall denote by $L(u, v)$ the number n of steps required in the algorithm. The question which must occur to many students when they first meet this algorithm is: How does $L(u, v)$—the work involved—depend on the sizes of u and v?

For a long time the following bounds on $L(u, v)$ have been known. Trivially $L(u, v) \geqq 1$, and $L(u, v) = 1$ if and only if $u \mid v$. We obtain an upper bound as follows. For given n, v is clearly least if $r_n = 1$ and each $q_i = 1$. In this case the equations (1) define u and v, respectively, as the nth and $(n+1)$st terms in the well-known Fibonacci sequence. It is known, and easily proved by induction, that the $(n+1)$st Fibonacci number is

$$f_{n+1} = \frac{\alpha^{n+1} + (-1)^n \alpha^{-n-1}}{\alpha + \alpha^{-1}},$$

where $\alpha = \frac{1}{2}(1 + \sqrt{5})$ is the positive root of the polynomial $X^2 - X - 1$. Since $\alpha + \alpha^{-1} = \sqrt{5}$ and $\alpha - \alpha^{-1} = 1$, a simple computation then shows $f_{n+1} \geqq 5^{-1/2}\alpha^n$. Thus we conclude that if $L(u, v) = n$, then $v \geqq f_{n+1}$ and so

$$(2) \qquad L(u, v) \leqq (\log v + 1)/\log \alpha = (2.078 \cdots)(\log v + 1).$$

In a recent paper [1] I have shown that this second bound is essentially of the right order. Indeed, if we put $\beta = 12\pi^{-2} \log 2 = 0.84276 \cdots$, then I prove that for each $\epsilon > 0$ and all $x > 1$

$$\left| L(u, v) - \beta \log v \right| < (\log v)^{\frac{1}{2}+\epsilon}$$

for almost all pairs u, v with $1 \leqq u \leqq v \leqq x$; the number of exceptional pairs is at most $x^2 \exp\{-c_0(\log x)^{\epsilon/2}\}$, where $c_0 > 0$ depends only on ϵ. This result seems to lie rather deep. Using quite different methods, however, it is possible to prove a weaker statement which is still qualitatively correct and this proof is quite elementary.

THEOREM. *Consider the proportion of pairs u, v with $1 \leqq u \leqq v \leqq x$ for which*

$$0.5 \log v \leqq L(u, v) \leqq 2.08(\log v + 1).$$

This proportion tends to 1 *as $x \to \infty$; indeed, the proportion of exceptions tends to* 0 *at least as fast as $x^{-0.01}$.*

REMARK. From (2) we know that the upper bound holds for all pairs. What we show below is that the lower bound holds for all except at most $5x^{1.99}$ pairs from a total of $\frac{1}{2}x(x+1)$ (when x is an integer).

Proof. We introduce the function $L_n(x)$ to denote the number of pairs u, v with $1 \leqq u \leqq v \leqq x$ and $L(u, v) = n$. Since $\log v \leqq \log x$, each pair u, v with $1 \leqq u \leqq v \leqq x$ which does not satisfy $L(u, v) \geqq \frac{1}{2} \log v$ must be counted in some $L_n(x)$ with $n < \frac{1}{2} \log x$. Hence we shall have proved the theorem when we have shown that

$$S(x) = \sum_{n<1/2 \log x} L_n(x) \tag{3}$$

is at most $5x^{1.99}$. We shall do this by estimating the value of $L_n(x)$.

Examining the equations (1) we see that u and v are completely determined by the values of $q_1, \cdots, q_n$ and r_n, and that $v \geqq q_1 \cdots q_n r_n$. Thus $L_n(x)$ is at most equal to the number of $(n+1)$-tuples $(q_1, \cdots, q_n, r_n)$ of positive integers such that $q_1 \cdots q_n r_n \leqq x$. Hence, for any real $s > 0$,

$$L_n(x) \leqq \sum x^s (q_1 \cdots q_n r_n)^{-s}, \tag{4}$$

where the sum is over all $(n+1)$-tuples $(q_1, \cdots, q_n, r_n)$ with product $\leqq x$. We only increase this sum if we take it over *all* $(n+1)$-tuples of positive integers, and so we conclude from (4) that for all $s > 1$

$$L_n(x) \leqq x^s \sum_{q_1=1}^{\infty} \cdots \sum_{q_n=1}^{\infty} \sum_{r_n=1}^{\infty} q_1^{-s} \cdots q_n^{-s} r_n^{-s} = x^s \zeta(s)^{n+1} \tag{5}$$

where $\zeta(s) = \sum_{q=1}^{\infty} q^{-s}$ is the Riemann zeta function.

Now from (3) and (5) we have (for all $s > 1$)

$$S(x) \leqq x^s \sum_{n<1/2 \log x} \zeta(s)^{n+1} < x^s \frac{\zeta(s)^{1/2 \log x+2}}{\zeta(s) - 1} \tag{6}$$

since $\zeta(s) > 1$. It remains to choose s judiciously. In fact $s = 3/2$ is close to the optimal choice, and for this value (6) gives

$$S(x) < x^{3/2}(\zeta(3/2) - 1)^{-1} \zeta(3/2)^{1/2 \log x+2} = Cx^\theta,$$

where $\theta = 3/2 + \frac{1}{2} \log \zeta(3/2)$ and $C = \zeta(3/2)^2(\zeta(3/2) - 1)^{-1}$. A table of the values of $\zeta(s)$ (for example, see [2]) shows that $2.61 < \zeta(3/2) < 2.62$, and so $\theta < 1.99$ and $C < 5$. This completes the proof of the theorem.

Acknowledgment. This work was supported in part by the National Research Council of Canada, Grant No. A7171.

References

1. J. D. Dixon, The number of steps in the Euclidean algorithm, J. Number Theory, 2 (1970)
2. E. Janke and F. Emde, Tables of Functions, Dover (reprint), New York, 1945.

THE MASSIVITY OF THE SQUARE

J. S. BYRNES, U. S. Naval Research Laboratory

If X is a compact subset of a metric space with metric ρ we define the **massivity** of X to be the following sequence:

$$\{R_2(X), R_3(X), R_4(X) \cdots\}, \quad \text{where} \quad R_n(X) = \max \min \rho(p_i, p_j), \tag{1}$$

the minimum taken over all i, j with $1 \leqq i < j \leqq n$ and the maximum over all possible choices of n points $p_1, p_2, \cdots, p_n$ in X. Thus massivity is seen to be a generalization of the concept of the diameter of a set.

The result proven by Tóth in [2] can easily be seen to imply that for a compact planar set X with area A, the quantity $R_n(X)$ is asymptotic to $\beta A^{1/2} n^{-1/2}$ as $n \to \infty$, where $\beta = 2^{1/2} 3^{-1/4}$. The purpose of this note is to present a new proof of this result for the square which yields an estimate for the error term.

THEOREM. *Let X be a square with area A, where $0 < A < \infty$. Then*

$$R_n(X) = \beta A^{1/2} n^{-1/2} + O(n^{-1}) \quad \textit{as } n \to \infty.$$

Proof. For convenience we assume that $A = 1$ and place the corners of X at $(0, 0)$, $(0, 1)$, $(1, 0)$, and $(1, 1)$. In addition we set $\alpha = \beta n^{-1/2}$.

First we construct a hexagonal lattice on X by placing points at $(0, 0)$, $(0, \alpha)$, $(0, 2\alpha)$, $\cdots$, $(0, [1/\alpha]\alpha)$ and using these intervals of length α as bases for equilateral triangles. (Note: $[x]$, as usual, denotes the greatest integer in x.) We continue the construction in this manner so that finally we have $[2\alpha^{-1}3^{-1/2}]+1$ rows with $[1/\alpha]+1$ points of our lattice in at least half of the rows and $[\alpha^{-1} - \frac{1}{2}]+1$ points in the remaining rows. We now require the following lemma:

LEMMA 1. *Let $x > 0$. Then $[x]+[x-\frac{1}{2}]+2 > 2x$.*

Proof. Let $[x] = m$, so that $m \leqq x < m+1$.

CASE 1: $m \leqq x < m + \frac{1}{2}$. Then $m + (m-1) + 2 = 2m+1 > 2x$.

CASE 2: $m + \frac{1}{2} \leqq x < m+1$. Then $m + m + 2 = 2m + 2 > 2x$.

By setting $x = \alpha^{-1}$ in Lemma I we see that in our lattice the average number of points per row is certainly greater than α^{-1}. Since there are more than $2\alpha^{-1}3^{-1/2}$ rows this shows that our lattice contains more than $2\alpha^{-2}3^{-1/2} = n$ points.

Thus

$$R_n > \beta/\sqrt{n}. \tag{2}$$

To find an upper bound for R_n let us suppose that we have a configuration C of n points in our square for which R_n is achieved. (The existence of such a configuration is guaranteed by a standard compactness argument.) We first require the following crude estimate.

LEMMA 2. $R_n \leqq 4(\pi n)^{-1/2}$.

Proof. Draw n circles with radius $\frac{1}{2}R_n$ and centers at the points of C. The interiors of these circles must be disjoint, and at least one quarter of the area of each of them must lie within our unit square. Thus, $\frac{1}{4}n\pi(\frac{1}{2}R_n)^2 \leqq 1$, and the lemma is proved.

We now let T be the (closed) square formed by removing a strip of width $\frac{1}{2}R_n\sqrt{3}$ from each side of X, and we let k be the number of points of C which are removed in this manner. Clearly these k points can be placed in a strip S of length 4 and width $\frac{1}{2}R_n\sqrt{3}$ without decreasing the distances between them (where S is open along one length and closed on the other three sides). Thus, if m is the maximum number of points that can be placed in S so that the distance between any two points is at least R_n, it follows that $k \leqq m$. If we divide S into $[8/R_n]$ substrips of length $\frac{1}{2}R_n$ (with possibly a piece of length less than $\frac{1}{2}R_n$ remaining) we see that at most one of the m points can be in each substrip, so that $m \leqq [8/R_n]+1 \leqq 8/R_n+1$. Thus at least $n-8/R_n-1$ points of C are in T. By (2) we have

(3) *At least* $n - 4(2n\sqrt{3})^{1/2} - 1$ *points of* C *are in* T.

If we now let K be the maximum number of disjoint (except for boundaries) circles of radius $\frac{1}{2}R_n$ that can be contained in the (closed) square formed by removing a strip of width $\frac{1}{2}R_n(\sqrt{3}-1)$ from each side of X, then (3) shows that

$$K \geqq n - 4(2n\sqrt{3})^{1/2} - 1. \tag{4}$$

By Lemma 2 and the circle packing property [3, pp. 35–37], these K circles will have their centers on a hexagonal lattice of side R_n. Thus, each one will take up an area inside our unit square of

$$\pi(\tfrac{1}{2}R_n)^2 + 2R_n^2(\tfrac{1}{4}\sqrt{3} - \tfrac{1}{8}\pi) = \tfrac{1}{2}R_n^2\sqrt{3}$$

(since the area of the region between three of these "ideally packed" circles is $R_n^2(\frac{1}{4}\sqrt{3}-\frac{1}{8}\pi)$).

Combining (4) with this we get $\frac{1}{2}[n-(2n\sqrt{3})^{1/2}-1]R_n^2\sqrt{3} \leqq 1$ which, with (2), yields

$$\frac{\beta}{\sqrt{n}} < R_n < \frac{\beta}{\sqrt{n}} + \frac{4}{n} + \frac{30}{n^{3/2}},$$

the desired result.

REMARKS. If we attempt to determine R_n of the unit square explicitly we find that even for small n the problem is decidedly nontrivial. The cases $n=2$, 3, and 5 are clear while the proof of the "obvious" solution for $n=4$ (one point in each corner) surprisingly presents some minor difficulties. Diagrams giving the correct configurations for $2 \leqq n \leqq 9$ appear in Schaer [4], where he presents a proof for $n=9$ and states that the case $n=7$ can be solved using essentially the same methods. For $n=6$ Schaer credits the solution to R. L. Graham, while for $n=8$ the solution was obtained independently by Vegh [6] and Schaer and Meir [5]. For $n>9$, however, it remains an interesting unsolved problem.

We might also note that the theorem holds if "square" is replaced by "convex set in the plane, or a finite union of such sets." The proof, which follows closely that given above, is left as an exercise.

In conclusion we observe that this problem is of far greater difficulty in Euclidean n space, $n>2$. In fact, the problem of the densest packing of spheres in three dimensions remains unsolved [1, p. 295, conjecture 1], so that even an asymptotic result for the massivity of the cube is unknown.

NAS-NRC Postdoctoral Research Associate, currently at the University of Massachusetts at Boston.

References

1. L. Fejes Tóth, Regular Figures, Macmillan, New York, 1964.

2. ———, Über einen geometrischen Satz, Math. Zeitschrift, 46 (1940) 83–85.

3. D. Hilbert, and S. Cohn-Vossen, Geometry and the Imagination, Chelsea, New York, 1952.

4. J. Schaer, The densest packing of 9 circles in a square, Canad. Math. Bull. no. 3, 8 (1965) 273–277.

5. J. Schaer and A. Meir, On a Geometric Extremum Problem, Canad. Math. Bull. no. 1, 8 (1965) 21–27.

6. E. Vegh, The Distances Between the Roots of an Algebraic Equation and the Conjecture by Leo Moser. Doctoral Thesis, University of North Carolina, 1965.

ON GENERATING PYTHAGOREAN TRIPLES

M. G. TEIGEN AND D. W. HADWIN, Augustana College

1. Introduction. For positive integers, a, b, c we shall call the ordered triple (a, b, c) a **Pythagorean triple** provided $a^2+b^2=c^2$. A Pythagorean triple (a, b, c) is called **primitive** if a and b are relatively prime. It is well known [1, pp. 63–66] that a primitive Pythagorean triple (a, b, c) in which a is even has a representation $a=2uv$, $b=u^2-v^2$, $c=u^2+v^2$, where u and v are positive integers. However, it is not possible to represent all Pythagorean triples in this way, e.g., (12, 9, 15).

The purpose of this paper is to present a new way of representing Pythagorean triples. This representation is very easy to understand and gives a simple way of generating all Pythagorean triples.

2. The representation.

THEOREM 1. *Suppose (a, b, c) is a Pythagorean triple. If x, y, z are defined by*

$$x = c - b, \qquad y = c - a, \qquad z = a + b - c, \tag{1}$$

then a, b, c may be represented in the form

$$a = x + z, \qquad b = y + z, \qquad c = x + y + z, \tag{2}$$

where x, y, z satisfy

$$x, y, z \text{ are positive}, \quad 2xy = z^2, \quad z \text{ is even}. \tag{3}$$

Conversely, if x, y, z are integers satisfying (3) *and a, b, c are defined by* (2), *then (a, b, c) is a Pythagorean triple satisfying* (1).

The proof of the above theorem requires only simple algebraic computations and is therefore omitted.

Theorem 1 tells us that we may generate any Pythagorean triple by choosing an even integer z, an integer x which is a factor of $z^2/2$, and letting $y = z^2/2x$.

It follows from (1) that the representation of the triple (a, b, c) in terms of x, y, z is unique. Thus each Pythagorean triple is generated exactly once by this method.

It follows from $b-a=y-x$ that $a<b$ if and only if $x<y$. Thus we may generate a Pythagorean triple (a, b, c) satisfying $a<b<c$ by choosing $x<y$.

3. Primitive triples. The following theorem tells how primitive Pythagorean triples may be generated.

THEOREM 2. *Let a, b, c, x, y, z be as in Theorem* 1. *Then a and b are relatively prime if and only if x and y are relatively prime.*

Proof. Suppose a and b are relatively prime. Since $2xy = z^2$, any factor common to x and y is necessarily a factor of z. From $a = z + x$ and $b = z + y$, it follows that any factor common to x and y must also be a common factor of a and b. Since a and b are relatively prime, x and y must also be relatively prime.

Conversely suppose that x and y are relatively prime. It follows from $2xy = z^2$ that there are integers s and t such that $z = 2st$ with $2s$ and t relatively prime, and (without loss of generality) $x = 2s^2$, $y = t^2$. Thus $a = x + z = 2s^2 + 2st = 2s(s+t)$ and $b = y + z = t^2 + 2st = t(t+2s)$. To prove that a and b are relatively prime it suffices to show that each of the factors $2s$, $s+t$ of a is relatively prime to the factors t, $t+2s$ of b. We shall show that $s+t$ and $2s+t$ are relatively prime. It follows from $2s = 2(t+2s) - 2(s+t)$ and $t = 2(s+t) - (t+2s)$ that any factor common to $s+t$ and $2s+t$ also must be a common factor of $2s$ and t. Since $2s$ and t are relatively prime, $s+t$ and $t+2s$ must be relatively prime.

4. Examples. If we wish to generate the Pythagorean triple (3, 4, 5) we let $x = 1$, $y = 2$, and $z = 2$. On the other hand if we let $z = 12$, then we must choose x and y as factors of 72. If we let $x = 36$ and $y = 2$, we obtain the Pythagorean triple (48, 14, 50).

Reference

1. N. H. McCoy, Introduction to Modern Algebra, Allyn and Bacon, Boston, 1960.

ON FENCHEL'S THEOREM

R. A. HORN, Johns Hopkins University

Fenchel's well-known theorem on the total curvature μ of a closed C^2 space curve consists of the inequality $\mu \geqq 2\pi$, together with the statement that equality holds if and only if the curve is a plane convex curve. Our purpose is to deduce this result with a symmetry argument that seems simpler than any proof appearing in the textbooks or recent literature.

Let $G = \{\boldsymbol{R}(s) \mid 0 \leqq s \leqq L\}$ be a closed C^2 space curve with arclength parameter s, unit tangent $\boldsymbol{T}(s) = (T_1, T_2, T_3) \equiv (d/ds)\boldsymbol{R}(s)$, and curvature

$$|\kappa(s)| \equiv \|(d/ds)\boldsymbol{T}(s)\|.$$

The ***tangent indicatrix*** of G is the curve $\Gamma = \{\boldsymbol{T}(s) \mid 0 \leqq s \leqq L\}$ on the unit sphere S, and the ***total curvature*** of G is the quantity $\mu \equiv \int_0^L |\kappa(s)| ds$. Clearly, the total curvature of G is just the length of its tangent indicatrix. Consider the following two lemmata:

LEMMA 1. *The tangent indicatrix Γ of a closed C^1 space curve G is not contained in any open hemisphere of S. It is contained in a closed hemisphere if and only if G is a plane curve.*

Proof: If Γ were contained in a hemisphere of S, we could perform a rotation to place it in the northern hemisphere. Thus, we may assume $T_3(s) \geqq 0$ for all $s \in [0, L]$ and since G is a closed curve we know that

$$0 = (\boldsymbol{R}(L) - \boldsymbol{R}(0))_3 = \left(\int_0^L \boldsymbol{T}(s)ds\right)_3 = \int_0^L T_3(s)ds = 0.$$

This shows that $T_3(s)$ cannot be strictly positive, and hence Γ cannot lie in an open hemisphere. Furthermore, since $T_3(s)$ is nonnegative it must vanish identically, i.e., $0 \equiv T_3(s) \equiv (d/ds)R_3(s)$, and hence G must lie in a plane $R_3 = \text{constant}$. Conversely, if G is a plane curve, then Γ lies on a great circle and hence is contained in a closed hemisphere.

LEMMA 2. *Let Γ be a closed rectifiable curve on S. If the length of Γ is less than 2π, then it is contained in an open hemisphere of S; if Γ has length 2π, then it is contained in a closed hemisphere of S.*

Proof. Let P be any point on Γ and let Q be the point of Γ such that the curve segments $\Gamma_1 = PQ$ and $\Gamma_2 = QP$ have equal length, $\Gamma = \Gamma_1 + \Gamma_2$. Rotate S so that P and Q are located symmetrically with respect to the north pole N, i.e., so that either $P = Q = N$ or so that P and Q have the same latitude but have longitudes differing by 180°. If Γ does not now intersect the equator, the conclusions follow. If Γ_1 intersects the equator at some point, construct the unique curve Γ_2' which is symmetric to Γ_1 with respect to N. Then Γ_2' has the same length as Γ_1, the closed curve $\Gamma' \equiv \Gamma_1 + \Gamma_2'$ has the same length as Γ, and there is

a pair of antipodal equatorial points on Γ'. But if we join these points with great semicircles (which are the geodesics on S), we see that if Γ_1 intersects the equator, then Γ' (and hence Γ) has length at least 2π, and that if Γ_1 crosses the equator into the open southern hemisphere, then Γ' must be strictly longer than 2π.

Thus, if Γ has length less than 2π, then Γ_1 cannot intersect the equator and if the length of Γ is exactly 2π, then Γ_1 cannot cross the equator. Since the same argument applies to Γ_2 we conclude that if Γ has length less than 2π, then it is contained in the open northern hemisphere and that if its length is exactly 2π, then it must lie in the closed northern hemisphere.

Fenchel's inequality now follows immediately from these two results: Since the tangent indicatrix of a closed C^2 space curve cannot lie in an open hemisphere of S its length must be at least 2π. If the length is exactly 2π, then the tangent indicatrix lies in a closed hemisphere of S and hence the original curve must be a plane curve. Using the notion of the *rotation index* as in [1], one can now complete the full proof of Fenchel's Theorem by showing that a plane curve is convex if and only if its tangent indicatrix has length 2π.

REMARKS: This proof grew out of a discussion with Laird E. Taylor about a problem posed by Robert Osserman. It is a pleasure to acknowledge their helpful suggestions as well as those of S. S. Chern, H. Flanders, and the referees.

Contrary to our original belief, this treatment is not the first completely elementary proof of Fenchel's inequality. The referees have remarked that Anthony Morse, A. S. Besicovitch, and H. Flanders have presented elementary proofs in their lectures, and that shortly after Fenchel's original publication [2] in 1929, H. Liebmann [6] published a similar elementary proof which has been almost totally forgotten. In a 1951 article, Fenchel [3] gives references to several different, less elementary, proofs, while recent textbooks such as [4] and [5] have followed a surface theory approach, introduced by K. Voss [7] in 1955. It is interesting to compare these proofs with ours and with each other; Fenchel's elegant pearl has certainly inspired a variety of settings.

This work was partially supported by the National Science Foundation under grant GP-13258.

References

1. S. S. Chern, ed., Studies in Global Geometry and Analysis, vol. 4, MAA Studies in Mathematics, 1967, 32.

2. W. Fenchel, Über Krümmung und Windung geschlossener Raumkurven, Math. Annalen, 101 (1929) 238–252.

3. ———, The differential geometry of closed space curves, Bull. Amer. Math. Soc., 57 (1951) 44–54.

4. H. W. Guggenheimer, Differential Geometry, McGraw-Hill, New York, 1963, 251–252.

5. D. Laugwitz, Differential and Riemannian Geometry, Academic Press, New York, 1965, 203–204.

6. H. Liebmann, Elementarer Beweis des Fenchelschen Satzes über die Krümmung geschlossener Raumkurven, S.-B. Preuss. Akad. Wiss. (Phys.-Math. Klasse) (1929) 392-393.

7. K. Voss, Eine Bemerkung über die Totalkrümmung geschlossener Raumkurven, Arch. Math., 6 (1955) 259–263.

ON A CHARACTERIZATION OF THE 2-SPHERE

KRISHNA AMUR, Karnatak University, Mysore State, India

1. Introduction. The purpose of this note is to give a new proof of the following known

THEOREM. *If Σ is a smooth, closed, convex, orientable surface in E^3 with the property that at each point the sum of the radii of principal curvatures is constant, that is*

$$2H/K = \text{constant},$$

where H is the mean curvature and K is the Gaussian curvature, then Σ is a sphere.

This is a special case of Liebmann-Süss theorem [1]: *A convex hypersurface in E^{n+1} is an n-sphere if one of the elementary symmetric functions of the principal radii of curvature is constant.* It also follows as a particular case of a theorem of Chern [2] on a characterization of the n-sphere in E^{n+1}.

In establishing the theorem, we use a parallel surface Σ' whose points are at a constant distance $2H/K$ along the normals to Σ. It turns out that Σ' is diffeomorphic to Σ and that $H'=-H$ and $K'=K$; consequently Σ' also has the properties of Σ. While Σ' may not be imbedded in E^3, it serves as a useful tool for obtaining information about Σ.

2. Preliminaries. If n denotes the unit normal to Σ at a point $\mathbf{x}$, the corresponding point of Σ' is given by

$$\mathbf{x}' = \mathbf{x} - an, \tag{2.1}$$

where $a=2H/K=\text{constant}$. We have

$$d\mathbf{x}' = d\mathbf{x} - a\,dn. \tag{2.2}$$

Choosing an orthonormal frame $\mathbf{e}_1$, $\mathbf{e}_2$ in the tangent space to Σ at $\mathbf{x}$ such that $\det(\mathbf{e}_1, \mathbf{e}_2, n)=+1$, we have

$$d\mathbf{x} = \sigma_1\mathbf{e}_1 + \sigma_2\mathbf{e}_2 \quad \text{and} \quad dn = \omega_1\mathbf{e}_1 + \omega_2\mathbf{e}_2, \tag{2.3}$$

where σ_i and ω_i are 1-forms. From (2.2) we have

$$d\mathbf{x}' = (\sigma_1 - a\omega_1)\mathbf{e}_1 + (\sigma_2 + a\omega_2)\mathbf{e}_2, \tag{2.4}$$

which shows that we can choose the orthonormal frame $\mathbf{e}_1$, $\mathbf{e}_2$ in the tangent space to Σ' at $\mathbf{x}'$, and that the normal to Σ' at $\mathbf{x}'$ is n. Denoting the 1-form coefficients of $\mathbf{e}_1$ and $\mathbf{e}_2$ in (2.4) by σ_1' and σ_2' respectively, and the area elements of Σ and Σ' by σ and σ' respectively, we have

$$\sigma' = \sigma_1' \wedge \sigma_2' = (1 - 2aH + a^2K)\sigma_1 \wedge \sigma_2 = \sigma_1 \wedge \sigma_2 = \sigma. \tag{2.5}$$

It follows from (2.5) that the rank of the differential of the map $\mathbf{x}\to\mathbf{x}'$ is 2 everywhere. Also from (2.1) it is clear that the map $\mathbf{x}\to\mathbf{x}'$ is bijective. Hence Σ' is diffeomorphic to Σ.

Since Σ and Σ' have the same normal at the corresponding points,

$$K'\sigma' = \omega_1' \wedge \omega_2' = \omega_1 \wedge \omega_2 = K\sigma, \tag{2.6}$$

which in view of (2.5) implies

$$K = K'. \tag{2.7}$$

Since from [3],

$$dn\bullet(dx \times n) = 2H\sigma \quad \text{and} \quad dn\bullet(dn \times n) = 2K\sigma, \tag{2.8}$$

we have

$$\begin{aligned} 2H'\sigma' &= dn\bullet(dx' \times n) \\ &= dn\bullet[(dx - a\,dn) \times n] \\ &= 2H\sigma - \frac{2H}{K}(2K\sigma) \\ &= -\,2H\sigma. \end{aligned}$$

Hence from (2.5) we get

$$H' = -\,H. \tag{2.9}$$

3. Proof of the theorem. Since Σ' is diffeomorphic to Σ, it is smooth, closed, and orientable. We can therefore apply the Minkowski formula

$$\int_\Sigma \sigma = \int_\Sigma Hp\sigma \tag{3.1}$$

to Σ' to obtain

$$\int_\Sigma \sigma = \int_\Sigma \sigma' = \int_\Sigma H'p'\sigma' = -\int_\Sigma H(p-a)\sigma = -\int_\Sigma Hp\sigma + 2\int_\Sigma \frac{H^2}{K}\sigma, \tag{3.2}$$

where we have used (2.1), (2.5), and (2.9). Using (3.1) again in (3.2), we get

$$\int_\Sigma \left(\frac{H^2}{K} - 1\right)\sigma = 0. \tag{3.3}$$

Clearly $H^2 - K \geqq 0$ holds, with equality at umbilical points, and since we have assumed that Σ is convex, $K > 0$. The integrand in (3.3) is therefore nonnegative, so $H^2 = K$ at all points of Σ. Hence the surface Σ must be a sphere.

References

1. T. Bonnesen and W. Fenchel, Theorie der konvexen Körper, Springer, Berlin, 1934.
2. S. S. Chern, Integral formulas for hypersurfaces in Euclidean space and their applications to uniqueness theorems, J. Math. Mech., 8 (1959) 947–955.
3. H. Flanders, Differential Forms, Academic Press, New York, 1963, pp. 40–44.

and so

$$Z^2 < 1 + \frac{2}{\mu^2} \cdot \tag{6}$$

Since X, Y, Z are positive integers, inequality (6) implies that Z can take only a finite number of values. The same holds for Y (on using (5)) and X (on using (3) or the relation $X < Y+Z$).

Since $X+Y+Z=\mu A$, A takes only a finite number of values. So do the sides a, b, c, in view of relations (2). Thus $N(\lambda)$ is finite for each positive $\lambda > 2$.

If $\lambda \leqq 2$ (i.e., $\mu \leqq 1$), we obtain the same conclusion on using the second inequality in (4) and proceeding as before.

We shall next show that $N(\lambda)=0$ for $\lambda > \sqrt{8}$ (i.e., $\mu > \sqrt{2}$). For such a value of μ we have from (6) that $Z=1$. The relation (5) then shows that $Y=1$. Since $X < Y+Z$, we now have $X=1$. Using relations (3) and (2), one obtains $\mu=3$, $A=\sqrt{3}$, $a=b=c=2$.

REMARKS. The remarks made at the beginning of this note show that $N(1)=5$ and $N(2)=1$. Our theorem shows that $N(3)=N(4)= \cdots =0$.

It would be interesting to consider a similar problem for a quadrilateral, and in general, a polygon of n sides ($n > 2$).

RESEARCH PROBLEMS

EDITED BY RICHARD GUY

In this Department the Monthly presents easily stated research problems dealing with notions ordinarily encountered in undergraduate mathematics. Each problem should be accompanied by relevant references (if any are known to the author) and by a brief description of known partial results. Material should be sent to Richard Guy, Department of Mathematics, Statistics, and Computing Science, The University of Calgary, Calgary 44, Alberta, Canada.

HOW OFTEN DOES AN INTEGER OCCUR AS A BINOMIAL COEFFICIENT?

DAVID SINGMASTER, Polytechnic of the South Bank, London

Let $N(a)$ be the number of times a occurs as a binomial coefficient, $\binom{n}{k}$. We have $N(1)=\infty$, $N(2)=1$, $N(3)=N(4)=N(5)=2$, $N(6)=3$, etc. Clearly, for $a>1$, $N(a)<\infty$. Below we establish that $N(a)=O(\log a)$. We conjecture that $N(a)=O(1)$, that is, that the number of solutions of $\binom{n}{k}=a$ is bounded for $a>1$. Erdös, in a private communication, concurs in this conjecture and states that it must be very hard. In a later communication, he suggests trying to show $N(a) = O(\log \log a)$.

If we let $M(k)$ be the first integer a such that $N(a)=k$, we have: $M(1)=2$, $M(2)=3$, $M(3)=6$, $M(4)=10$, $M(6)=120$. The next values would be interesting to know.

PROPOSITION. $N(a)=O(\log a)$.

Proof: Let b be the first b such that $\binom{2b}{b}>a$. Now

$$\binom{i+j}{i}=\binom{i+j}{j}$$

is monotonically increasing in i and in j; hence

$$\binom{b+i+b+j}{b+i}\geqq\binom{b+b+j}{b}\geqq\binom{2b}{b}>a \qquad \text{for all } i,j\geqq 0.$$

Thus $\binom{i+j}{j}=a$ implies $i<b$ or $j<b$. Again by monotonicity, for each value of i (or j),

$$\binom{i+j}{j}=a$$

has at most one solution. Hence $N(a)<2b$. Now $\binom{2b}{b}\geqq 2^b$, so we have

$$a\geqq\binom{2(b-1)}{b-1}\geqq 2^{b-1};$$

hence $b\leqq\log_2 a+1$, and $N(a)\leqq 2+2\log_2 a=O(\log a)$.

Added in Proof: I have recently found that $M(8)=3003$ and this is the only solution to $N(a)\geqq 8$ with $a\leqq 2^{23}$. There are six solutions to $N(a)=6$ with $a\leqq 2^{23}$, namely: 120, 210, 1540, 7140, 11628, and 24310.

HOW IS A GRAPH'S BETTI NUMBER RELATED TO ITS GENUS?

R. A. DUKE, University of Washington

If a connected graph G has v vertices and e edges, the number $1-v+e$ is called the (1-dimensional) **Betti number** of G, denoted by $\beta(G)$. This value, which is nonnegative for connected G, was one of the first numerical characteristics of a graph studied, having been introduced by von Staudt [11] and Kirchhoff [5]. The Betti number $\beta(G)$ is also the rank of the fundamental group of G (see [9], Ch. 6) and is related to the **Euler characteristic** of G, $\chi(G)$, by the equation $\beta(G)=1-\chi(G)$.

The genus of a graph is defined in terms of its embeddings in compact, closed, orientable 2-dimensional manifolds. Each such manifold can be formed by attaching an appropriate number of "handles" to a copy of the 2-sphere. For a manifold M, the number of handles required is called the **genus** of M, denoted by $\gamma(M)$, and is related to the Euler characteristic of M by the formula $2\gamma(M)=2-\chi(M)$. Each graph can be embedded in some orientable 2-manifold by placing its vertices on the 2-sphere and attaching one handle to accommodate each edge. In general, however, such a construction yields a manifold of unneces-

sarily large genus. The **genus** $\gamma(G)$ of the graph G is the smallest of the numbers $\gamma(N)$ for orientable 2-manifolds N in which G can be embedded.

While it can be shown that the Betti number of a connected graph is greater than twice its genus, the exact relationship between these numbers is unknown, and it appears that no counterexample has been found to the following conjecture of [2].

CONJECTURE. *If G is a connected graph, then* $\beta(G) \geqq 4\gamma(G)$.

That the conjectured inequality does hold if $\gamma(G)=1$ is implied by the following results.

A graph is said to be **nonplanar** provided that it is of positive genus, that is, cannot be embedded in the 2-sphere (and hence not in the plane). The first characterization of nonplanar graphs was provided by Kuratowski [8] who showed that a graph G has genus zero if and only if G does not contain a subgraph homeomorphic to either the complete 5-graph K_5 or the 3×3 bipartite graph $K_{3,3}$, where K_5 has five vertices, each pair joined by an edge, and $K_{3,3}$ has vertices a_1, a_2, a_3 and b_1, b_2, b_3, with an edge joining a_i to b_j for each i and j, $1 \leqq i, j \leqq 3$. (For a proof see [1].)

The Kuratowski graphs K_5 and $K_{3,3}$ have Betti numbers 4 and 6 respectively. Thus, since the Betti number of a graph is at least as large as that of any of its subgraphs, it follows that for each nonplanar graph we have $\beta(G) \geqq 4$.

The graphs K_5 and $K_{3,3}$ each have the property that the deletion of any one of their edges yields a subgraph of lower genus. Such graphs are called **critical.** Thus K_5 and $K_{3,3}$ are the only critical graphs of genus one. Since the deletion of edges from a graph must eventually lower its genus, each graph of genus greater than or equal to n contains a critical graph of genus n as a subgraph.

For each positive integer n, one can construct a critical graph G of genus n having $\beta(G)=4n$ by joining n copies of the graph K_5 at a single point. Other examples of critical graphs with $\gamma(G)=2$ and $\beta(G)=8$ can be found in [4]. A complete listing of the critical graphs of genus n would provide a characterization of graphs of genus less than n, and it would suffice to establish the conjecture for the set of all critical graphs of genus n for each positive integer n. However, over 60 nonhomeomorphic critical graphs of genus two are known [3] and no proof that the collection of all such graphs is even finite has ever been published.

An embedding of a graph G in a 2-manifold M is called a **cellular embedding** if each component of the complement $M\backslash G$ is homeomorphic to an open disc in the plane. For a cellular embedding, the relationship between $\chi(G)$ and $\chi(M)$ yields the following equation

$$c = 1 + \beta(G) - 2\gamma(M), \tag{1}$$

where c is the number of components of $M\backslash G$. It was first shown by König [6, 7] that an embedding of G in an orientable 2-manifold of genus $\gamma(M)=\gamma(G)$ is cellular, and that therefore equation (1) can be replaced by

$$d(G) = 1 + \beta(G) - 2\gamma(G), \tag{2}$$

where $d(G)$ is the largest possible number of components of $M\backslash G$ for any embedding of G in some orientable 2-manifold M. (See also [15].) Since $d(G) \geqq 1$, we have immediately that $\beta(G) \geqq 2\gamma(G)$. In [2] it was shown that $\beta(G) \geqq 2\gamma(G)+2$. It also follows from (2) that the relationship stated in the conjecture is equivalent to the condition that $d(G) \geqq 2\gamma(G)+1$ for each connected graph G and also equivalent to $2d(G) \geqq \beta(G)+2$.

Another direct consequence of (1) is that if there exists any cellular embedding of a graph G in an orientable 2-manifold M, then $\gamma(M) \leqq \frac{1}{2}\beta(G)$. Recently it was proved that there do exist cellular embeddings of G in the orientable 2-manifold whose genus is the largest integer less than or equal to $\frac{1}{2}\beta(G)$ in the cases where G is the complete n-graph K_n [10] or the $m \times n$ bipartite graph $K_{m,n}$ [12]. In [2] it was shown that if there exist cellular embeddings of a connected graph G in orientable 2-manifolds of genera r and s, then there exists a cellular embedding of G in the orientable 2-manifold of genus k for each integer k satisfying $r \leqq k \leqq s$. Thus, since $\gamma(K_n)$ and $\gamma(K_{m,n})$ are known for all positive integers m and n ([14] and [13] respectively), the complete range of values $\gamma(M)$ for M in which G has a cellular embedding is known when G is one of these special graphs.

References

1. G. A. Dirac and S. Schuster, A theorem of Kuratowski, Nederl. Akad. Wetensch. Proc. Ser. A, 57 (1954) 343–348.

2. R. A. Duke, The genus, regional number, and Betti number of a graph, Canad. J. Math., 18 (1966) 817–822.

3. G. Haggard, Embeddings of graphs in surfaces, Ph.D. thesis, Purdue University, 1968.

4. I. N. Kagno, The mapping of graphs on surfaces, J. Math. Phys., 16 (1937) 46–75.

5. G. Kirchhoff, Über die Auflösung der Gleichungen auf welche man bei Untersuchung der linearen Verteilung galvanischer Ströme geführt wird, Ann. Physik., 72 (1847) 497–508.

6. D. König, Vonalrendszerek kétoldalú felületeken (Graphs on two-sided surfaces), Math. és természettud. Értesitö, 29 (1911) 112–117.

7. ———, A vonalrandszereck nemszámáról (On the genus of graphs), Math. és természettud. Értesitö, 29 (1911) 345–350.

8. K. Kuratowski, Sur le problème des courbes gauches en topologie, Fund. Math., 15 (1930) 271–283.

9. W. Massey, Algebraic Topology: An Introduction, Harcourt, Brace and World, New York, 1967.

10. E. Nordhaus, B. Stewart and A. White, On the maximum genus of a graph, J. Comb. Th., to appear.

11. G. K. Chr. v. Staudt, Geometrie der Lage, Nürnberg, 1847.

12. R. D. Ringeisen, The maximum genus of the complete bipartite graph $K_{m,n}$, to appear.

13. G. Ringel, Das Geschlecht des vollständiger paaren Graphen, Abh. Math. Sem. Univ. Hamburg, 28 (1965) 139–150.

14. ———, and J. W. T. Youngs, Solution of the Heawood map-coloring problem, Proc. Nat. Acad. Sci. USA, 60 (1968) 438–445.

15. J. W. T. Youngs, Minimal imbeddings and the genus of a graph, J. Math. Mech., 12 (1963) 303–315.

CLASSROOM NOTES

EDITED BY DAVID DRASIN

After May 1, 1971, manuscripts for this Department should be sent to Robert Gilmer, Department of Mathematics, Florida State University, Tallahassee, FL 32306.

THE KANTOROVICH THEOREM FOR NEWTON'S METHOD

R. A. TAPIA, Rice University

1. Introduction. Kantorovich's theorem on Newton's method is of fundamental importance and is bound to become standard fare in college courses on numerical methods; however, due to length and complexity of the earlier proofs, it has not been adequately presented as yet. The original proof [2] given by Kantorovich used recurrence relations. A nice treatment of this theorem using recurrence relations can be found in Rall [5]. In [3] Kantorovich and Akilov give a proof based on the concept of a majorant function. In a previous Classroom Note of this MONTHLY [4] Ortega gives a short and elegant proof using majorant functions. However, both the proofs using majorant functions, [3] and [4], fail to yield the important error estimates which Kantorovich's theorem is known to give; consequently the error estimates are left to be derived using the Kantorovich recurrence relations.

In this note we add to Ortega's proof a very short and simple derivation of the optimum error estimates which does not use the Kantorovich recurrence relations. The optimum error estimates are a natural consequence of our proof and are derived with less effort than the Kantorovich error estimates. Dennis [1] has derived these optimum error estimates in a lengthy but ingenious manner using the Kantorovich recurrence relations.

2. The Kantorovich Theorem. Kantorovich's theorem asserts that the iterative method of Newton, applied to a most general system of nonlinear equations $P(x)=0$, converges to a solution x^* near some given point x_0, provided the Jacobian of the system satisfies a Lipschitz condition near x_0 (see (iii) below) and its inverse at x_0 satisfies certain boundedness conditions (see (i) and (ii) below). The theorem also gives computable error bounds for the iterates.

The theorem and its proof take a very concise form when presented in the framework of normed linear spaces. The system of equations takes the form $P(x)=0$, where P is a general map from a Banach space X into a Banach space Y. The Jacobian of P at $x_0 \in X$ is the so-called Fréchet differential $P'(x_0)$, which is the linear map of X into Y defined by the relation

$$\lim_{\|\Delta x\| \to 0} \frac{\|P(x_0 + \Delta x) - P(x_0) - P'(x_0)(\Delta x)\|}{\|\Delta x\|} = 0,$$

see [5, p. 87]. It is assumed that P is defined and has a Fréchet differential at each point of a given convex open set $D_0 \subset X$.

THEOREM (Kantorovich): *Assume for some* $x_0 \in D_0$ *that* $[P'(x_0)]^{-1}$ *exists and that*

(i) $\quad \|[P'(x_0)]^{-1}\| \leqq B,$

(ii) $\quad \|[P'(x_0)]^{-1}P(x_0)\| \leqq \eta,$

(iii) $\quad \|P'(x) - P'(y)\| \leqq K\|x - y\|,$ *for all* x *and* y *in* D_0,

with $h = BK\eta \leqq \frac{1}{2}$.

$$Let\ \Omega_* = \{x \mid \|x - x_0\| \leqq t^*\}, \quad where\ t^* = \left(\frac{1 - \sqrt{1 - 2h}}{h}\right)\eta.$$

Now, if $\Omega_* \subset D_0$, *then the Newton iterates,* $x_{k+1} = x_k - [P'(x_k)]^{-1}P(x_k)$, *are well defined, remain in* Ω_*, *and converge to* $x^* \in \Omega_*$ *such that* $P(x^*) = 0$. *In addition*

$$\|x^* - x_k\| \leqq \frac{\eta}{h}\left(\frac{(1 - \sqrt{1 - 2h})^{2^k}}{2^k}\right), \qquad k = 0, 1, 2, \cdots.$$

Proof. (*Ortega*). Let

$$p(t) = \frac{h}{2\eta}t^2 - t + \eta, \quad s(t) = t - \frac{p(t)}{p'(t)}, \quad \text{and} \quad t_{k+1} = s(t_k)$$

with $t_0 = 0$. A straightforward calculation shows that for $k \geqq 1$

$$t_{k+1} - t_k = \frac{h}{2}\,\frac{(t_k - t_{k-1})^2}{(\eta - ht_k)}. \tag{1}$$

The mean value theorem [5, p. 122] gives $\|P(x) - P(y) - P'(y)(x - y)\| \leqq K/2\|x - y\|^2$ for all x and y in D_0. If x is in the interior of Ω_*, then

$$\|x - x_0\| < t^* \leqq \frac{1}{BK} \quad \text{so} \quad \|P'(x) - P'(x_0)\| < \frac{1}{B},$$

and by Banach's theorem [7, p. 164] $[P'(x)]^{-1}$ exists with

$$\|[P'(x)]^{-1}\| \leqq \frac{B}{1 - BK\|x - x_0\|}.$$

If also $N(x) = x - [P'(x)]^{-1}P(x)$ is in the interior of Ω_*, then since $P(x) + P'(x) \cdot (N(x) - x) = 0$,

$$\begin{aligned} \|N(N(x)) - N(x)\| &= \|P'(N(x))^{-1}P(N(x))\| \\ &\leqq \frac{B\|P(N(x)) - P(x) - P'(x)(N(x) - x)\|}{1 - BK\|x_0 - N(x)\|} \\ &\leqq \frac{h\|x - N(x)\|^2}{2(\eta - h\|x_0 - N(x)\|)}. \end{aligned} \tag{2}$$

Clearly $s'(t)>0$ for $0\leqq t<t^*$ so $t_k<t_{k+1}\to t^*$. Now if x_n exists and $\|x_n-x_{n-1}\|\leqq t_n-t_{n-1}$ for $n\leqq k$, as is true by assumption for $k=1$, then $\|x_k-x_0\|\leqq t_k-t_0<t^*$ so x_k is in the interior of Ω_*. Evaluating (2) at $x=x_{k-1}$ and using (1), we have $\|x_{k+1}-x_k\|\leqq t_{k+1}-t_k$ for all k. It follows that $\|x_{k+m}-x_k\|\leqq t^*-t_k$ for arbitrary positive integers k and m. Hence $\{x_k\}$ is a Cauchy sequence and must therefore converge to $x^*\in\Omega_*$ with $\|x^*-x_k\|\leqq t^*-t_k$. That $P(x^*)=0$ follows from the continuity of P and P' at x^*.

Multiplying (1) by h we see that if ht_k-ht_{k-1} and ht_k (as functions of h) increase with h, as is true for $k=1$, then $ht_{k+1}-ht_k$ and therefore ht_{k+1} increases with h. When $h=\frac{1}{2}$, $\eta=ht_k=\eta/2^k$ so $1/(\eta-ht_k)\leqq 2^k/\eta$. We now have

$$t^*-t_{k+1}=\frac{h(t^*-t_k)^2}{2(\eta-ht_k)}\leqq\frac{h}{\eta}2^{k-1}(t^*-t_k)^2.$$

Therefore

$$t^*-t_{k+1}\leqq\frac{\eta}{h}\frac{[\alpha(h)]^{2^{k+1}}}{2^{k+1}}\quad\text{whenever}\quad t^*-t_k\leqq\frac{\eta}{h}\frac{[\alpha(h)]^{2^k}}{2^k}.$$

Since $t^*-t_0=(\eta/h)(1-\sqrt{1-2h})$, it follows that $\alpha(h)=1-\sqrt{1-2h}$ gives the optimum error estimate of this form.

REMARK. Kantorovich's error estimates, $\|x^*-x_k\|\leqq(\eta/h)(2h)^{2^k}/2^k$, follow since $1-\sqrt{1-2h}\leqq 2h$ for $0\leqq h\leqq\frac{1}{2}$. A simple calculation shows that for $k=0, 1$ we actually have

$$t^*-t_k=\frac{\eta}{h}\frac{(1-\sqrt{1-2h})^{2^k}}{2^k};$$

consequently not only is this error estimate the best of the form

$$\|x^*-x_k\|\leqq\frac{\eta}{h}\frac{(\alpha(h))^{2^k}}{2^k},$$

but it is also extremely accurate for small k. Kantorovich's estimates give equality only when $h=\frac{1}{2}$. Equality would hold for $h=\frac{1}{2}$ using any error estimate of this form, since we clearly require $\alpha(h)\leqq 1$.

REMARK. If

$$t_{k+1}-t_k\leqq\frac{\eta}{h}\frac{[\alpha(h)]^{2^k}}{2^{k+1}},\tag{3}$$

then recalling that $1/(\eta-ht_k)\leqq 2^k/\eta$, we have from (1) that (3) holds for $k=n+1$ and therefore by induction for $k\geqq n$. Summing (3) we have $t^*-t_k\leqq(\eta/h)\cdot[\alpha(h)]^{2^k}/2^k$ for $k\geqq n$. Now if we wish to find the smallest $\alpha(h)$ such that (3) holds for $k\geqq 0$ we must have $t_1-t_0=\eta\leqq(\eta/h)\alpha(h)/2$ or $\alpha(h)\geqq 2h$; consequently the best error estimates of the form $\|x^*-x_k\|\leqq(\eta/h)[\alpha(h)]^{2^k}/2^k$ which hold for

$k \geqq 0$ and can be derived in this manner are the Kantorovich error estimates, i.e., $\alpha(h)=2h$. Even if $t_{k+1}-t_k \nleqq (\eta/h)[\alpha(h)]^{2^k}/2^{k+1}$, it may still be true that $t^*-t_k \leqq (\eta/h)[\alpha(h)]^{2^k}/2^k$; this, however, would require a different method of proof. In particular, the best estimate of this form such that $t_{k+1}-t_k \leqq (\eta/h) \cdot [\alpha(h)]^{2^k}/2^{k+1}$ for $k \geqq 1$ is $\alpha(h)=\sqrt{2h^2/(1-h)}$, which also has the property that $t^*-t_k \leqq (\eta/h)[\alpha(h)]^{2^k}/2^k$ for $k \geqq 0$. It is highly unlikely that this method of proof could lead us to the optimum error estimates where $\alpha(h)=1-\sqrt{1-2h}$ since a very tedious and unimaginative calculation shows at least for $0 \leqq k \leqq 5$, that $t_{k+1}-t_k \nleqq (\eta/h)(1-\sqrt{1-2h})^{2^k}/2^{k+1}$. Most methods of proof lead very naturally to the Kantorovich estimates. Our derivation is best appreciated by investigating these other methods.

Added in proof. In their recent book Ortega and Rheinboldt give a proof of the Kantorovich error estimates which does not use the Kantorovich recurrence relations. Their proof can also be used to obtain the optimum error estimates.

This work was sponsored by the Mathematics Research Center, United States Army, Madison, Wisconsin, under Contract No. DA-31-124-ARO-D-462 and was performed while the author was a member of the Mathematics Research Center.

References

1. J. E. Dennis, On the Kantorovich Hypothesis for Newton's method, SIAM J. Numer. Anal., 6 (1969) 493–507.

2. L. V. Kantorovich, Functional Analysis and Applied Mathematics, translated by C. D. Benster, National Bureau of Standards Report 1509, 1952.

3. L. V. Kantorovich and G. P. Akilov, Functional Analysis in Normed Spaces, Pergamon, New York, 1964.

4. J. M. Ortega, The Newton-Kantorovich Theorem, this MONTHLY, 75 (1968) 658–660.

5. L. B. Rall, Computational Solution of Nonlinear Operator Equations, Wiley, New York, 1969.

6. R. A. Tapia, The Kantorovich Theorem, Mathematics Research Center, Report #1007, University of Wisconsin, Madison, July 1969.

7. A. E. Taylor, Introduction to Functional Analysis, Wiley, New York, 1957.

ON EXTENSIONS OF Q BY SQUARE ROOTS

R. L. ROTH, University of Colorado

In courses on field theory one often uses the field of all algebraic numbers as an example of an extension of Q, the field of rational numbers, which is both algebraic and of infinite degree over Q. A more intuitively grasped example is perhaps the field $Q(\sqrt{2}, \sqrt{3}, \sqrt{5}, \cdots, \sqrt{p}, \cdots)$, where we adjoin the square roots of all the positive primes to Q. However, to prove that each new adjunction of the square root of a prime is in fact a proper extension, it turns out to be easier to prove the following stronger statement by induction. (In this note, "prime" will always mean positive prime integer.)

THEOREM. *Let $p_1, \cdots, p_n$ be n distinct positive prime integers. Let $F=Q(\sqrt{p_1}, \cdots, \sqrt{p_n})$. Let $q_1, \cdots, q_r$ be any set of distinct positive primes none of which appear in the list $\{p_1, \cdots, p_n\}$. Then $\sqrt{q_1q_2 \cdots q_r} \notin F$.*

Proof. By induction on n.

Let $n=0$. Then $F=Q$. If $q_1, \cdots, q_r$ are any distinct primes, then $x^2-q_1q_2 \cdots q_r$ is irreducible over Q by Eisenstein's criterion and $\sqrt{q_1q_2 \cdots q_r} \notin Q$. (This can also be shown by a proof analogous to the classical proof that $\sqrt{2} \notin Q$.)

Now assume the theorem is true for $n-1$ and suppose that

$$F = Q(\sqrt{p_1}, \cdots, \sqrt{p_{n-1}}, \sqrt{p_n}).$$

Set $F_0=Q(\sqrt{p_1}, \cdots, \sqrt{p_{n-1}})$. By induction the theorem holds for F_0, and $F=F_0(\sqrt{p_n})$ is an extension of F_0 of degree 2. Let $q_1, q_2, \cdots, q_r$ be any distinct primes not on the list $\{p_1, \cdots, p_n\}$. Suppose it were true that $\sqrt{q_1q_2 \cdots q_r} \in F$. Then $\sqrt{q_1q_2 \cdots q_r}=a+b\sqrt{p_n}$ where a and b belong to F_0. Then

$$q_1q_2 \cdots q_r = a^2 + b^2p_n + 2ab\sqrt{p_n}. \tag{1}$$

(i) If a and b are both not 0, equation (1) shows that

$$\sqrt{p_n} = \frac{q_1q_2 \cdots q_r - a^2 - p_nb^2}{2ab}$$

would lie in F_0.

(ii) If $b=0$, then $\sqrt{q_1 \cdots q_r}=a \in F_0$.

(iii) If $a=0$, then $\sqrt{q_1 \cdots q_r}=b\sqrt{p_n}$; hence $\sqrt{q_1 \cdots q_rp_n}=p_nb \in F_0$. However, each of these three possibilities contradicts the inductive hypothesis (with respect to F_0). So $\sqrt{q_1 \cdots q_r} \notin F$.

COROLLARY 1. *If q is not on the list of primes $\{p_1, \cdots, p_r\}$ and q is a prime, then $\sqrt{q} \notin Q(\sqrt{p_1}, \cdots, \sqrt{p_n})$.*

COROLLARY 2. *If $p_1, \cdots, p_n$ are any distinct primes then*

$$[Q(\sqrt{p_1}, \sqrt{p_2}, \cdots, \sqrt{p_n}):Q] = 2^n.$$

COROLLARY 3. *If $p_1, p_2, \cdots, p_i, \cdots$ is any infinite sequence of distinct primes then $Q(\sqrt{p_1}, \sqrt{p_2}, \cdots, \sqrt{p_i}, \cdots)$ is an infinite algebraic extension of Q. In particular $Q(\sqrt{2}, \sqrt{3}, \sqrt{5}, \cdots, \sqrt{p}, \cdots)$ where p runs over all primes has this property.*

(Note, the latter is the same as the field $Q(\sqrt{1}, \sqrt{2}, \sqrt{3}, \cdots, \sqrt{n}, \cdots)$ where n runs over all positive integers.)

The theorem may also be applied in special examples, for instance $[Q(\sqrt{2}, \sqrt{7}, \sqrt{15}):Q]=8$. $[Q(\sqrt{14}, \sqrt{15}):Q]=4$ since $\sqrt{15} \notin Q(\sqrt{14}) \subset Q(\sqrt{2}, \sqrt{7})$. $[Q(\sqrt{14}, \sqrt{6}):Q]=4$, since if $\sqrt{14} \in Q(\sqrt{6}) \subset Q(\sqrt{2}, \sqrt{3})$, then $\sqrt{7} \in Q(\sqrt{2}, \sqrt{3})$.

TWO NOTES ON HOMOLOGICAL ALGEBRA

G. W. McCollum, Harvard University

1. On the serpent lemma. This note contains a new version of the "serpent lemma." It takes the form of a refinement of the usual construction of the long exact homology sequence, and in particular contains the usual construction. It also implies the version of the serpent lemma in Bourbaki (Alg. Comm., Chapter 1, Section 1, No. 4, Prop. 2). However, it has the advantage of a simplicity and symmetry of statement which make it easier to remember.

Serpent Lemma. Consider a sequence of chain complexes and chain maps

$$
\begin{array}{ccccccccc}
 & & \vdots & & \vdots & & \vdots & & \\
 & & \downarrow & & \downarrow & & \downarrow & & \\
0 & \to & C'_{i+1} & \to & C_{i+1} & \to & C''_{i+1} & \to & 0 \\
 & & \downarrow & & \downarrow & & \downarrow & & \\
0 & \to & C'_{i} & \to & C_{i} & \to & C''_{i} & \to & 0 \\
 & & \downarrow & & \downarrow & & \downarrow & & \\
0 & \to & C'_{i-1} & \to & C_{i-1} & \to & C''_{i-1} & \to & 0 \\
 & & \downarrow & & \downarrow & & \downarrow & & \\
0 & \to & C'_{i-2} & \to & C_{i-2} & \to & C''_{i-2} & \to & 0 \\
 & & \downarrow & & \downarrow & & \downarrow & & \\
 & & \vdots & & \vdots & & \vdots & &
\end{array}
$$

(with no exact conditions except as stated below). If the diagram is exact at the points of the diagonal C'_{i-1}, C_i, C''_{i+1} (where exactness throughout means "horizontal" exactness), then the sequence

$$H'_i \to H_i \to H''_i$$

is exact, where the H's are the homology groups. If in addition the diagram is exact at the points of the diagonal C'_{i-2}, C_{i-1}, C''_i then the sequence

$$H'_i \to H_i \to H''_i \to H'_{i-1} \to H_{i-1} \to H''_{i-1}$$

is exact, where the map from H''_i to H'_{i-1} is the usual boundary map.

The proof is a routine diagram chase, identical to the usual one for the long exact homology sequence.

Bourbaki's "Proposition" cited above easily follows from this version of the serpent lemma by using a diagram in which most of the modules are zero. In fact, the result most naturally derived is slightly sharper than Bourbaki's, since exactness of (u', v') is not needed in part (i) of his Proposition, and exactness of (u, v) is not needed in part (ii).

2. Diagram Chasing. In homological algebra a map is often defined by composing functions and converses of functions. (The converse R^{-1} of a relation R is the set of all (x, y) such that $(y, x) \in R$.) Checking commutativity of diagrams involving such maps is often tiresome. The result in this paper gives a method of immediately checking commutativity in such situations.

PROPOSITION. *Consider the diagram*

$$\begin{array}{ccc} A & \overset{R}{\longleftrightarrow} & B \\ f\downarrow & & \downarrow g \\ A' & \underset{R'}{\longleftrightarrow} & B' \end{array}$$

where A, B, A', B' are sets, R and R' are relations, and f and g are functions. Then $gR \subset R'f$ if and only if $fR^{-1} \subset R'^{-1}g$.

This is easy to check; in fact, either inclusion is equivalent to

$$(a, b) \in R \Rightarrow (f(a), g(b)) \in R'.$$

In the applications it is important to note that one function $f: A \to B$ contains a function $g: A \to B$ if and only if they are equal.

Example. Consider the diagram of sets and functions

$$\begin{array}{ccccc} A_0 & \overset{\alpha}{\longleftarrow} & A_1 & \overset{\beta}{\longrightarrow} & A_2 \\ \downarrow f_0 & & \downarrow f_1 & & \downarrow f_2 \\ A_0' & \underset{}{\overset{\alpha'}{\longleftarrow}} & A_1' & \overset{\beta'}{\longrightarrow} & A_2' \end{array}$$

Assume that both squares commute and that $\beta\alpha^{-1}$ is a function and $\beta'\alpha'^{-1}$ is a function. Then $f_2\beta\alpha^{-1} \subset \beta' f_1 \alpha^{-1} \subset \beta'\alpha'^{-1} f_0$ by the proposition. But since the first and last members are functions, they are equal. So the diagram

$$\begin{array}{ccc} A_0 & \overset{\beta\alpha^{-1}}{\longrightarrow} & A_2 \\ f_0 \downarrow & & \downarrow f_2 \\ A_0' & \overset{\beta'\alpha'^{-1}}{\longrightarrow} & A_2' \end{array} \qquad \text{commutes.}$$

MATHEMATICAL EDUCATION

EDITED BY J. G. HARVEY AND M. W. POWNALL

Material for this Department should be sent to either of the editors: J. G. Harvey, Department of Mathematics, University of Wisconsin, Madison, WI 53706; M. W. Pownall, Department of Mathematics, Colgate University, Hamilton, NY 13346.

A COMPUTER-ASSISTED GRADING PROGRAM

D. R. FORBES, University of Saskatchewan, Regina Campus

If published reports are indicative of practice, the interest of mathematics teachers in the computer has been primarily to introduce it as a tool for problem solving. This use of the computer in elementary calculus courses has undoubtedly been advanced by the activities of the CUPM Panel on Computing.

Perhaps we could also be using the computer to greater advantage in assisting our classroom efforts at the precalculus level. This note briefly describes an experimental program we tried last semester in one class of introductory mathematics.

The class was a theatre section of a course intended for those who wish to take the calculus, but who need a semester of preparation and review. It consisted of a rapid review of high school material as well as new topics. Most students in the course averaged between 65 and 80 percent on their provincial high school final examinations in mathematics. For many students it satisfies their need of one course in logic or mathematics.

Each week every student in the class turned in an assignment and wrote a fifty minute quiz, both of which were prepared so that they could be computer graded. Construction of these assignments and quizzes was time-consuming as few research-backed or even well-thought-out materials are available. Task analysis is involved in generating the multiple choice responses and for the first few weeks a random sample of both quizzes and assignments were handwritten and graded as a control. That is, the same question was in multiple choice format for some and in written answer format for others. This indicated whether or not the multiple choice format affected student grades for a given week. Also, more than one quiz and assignment were designed for each week and then students were randomly assigned to quiz and assignment groups.

The file of these assignments and quizzes can now be used to generate new assignments and quizzes that are entirely based on demonstrated student frequency response rather than instructor expectation of error responses.

The true-false or obvious multiple guess paradigms can be avoided after experience if the test generator is creative. A library of student responses is essential for reliable construction of computer graded materials. As many departments keep examinations on file for several years, these are often available.

Each student was assigned a weekly composite grade which was a weighted average of his performance on both the quiz and assignment. If this was a fail-

ing grade he was scheduled for an interview with the grading assistant. As the department has records of student class schedules, this interview assignment was automatically done. If a student received a non-failing grade, he was simply given his mark. All students were free to see the instructor during regularly scheduled office hours.

The interview was conducted by a grader. Such graders are assigned to all theatre sections so the project involved no additional staff commitments.

The interviewer had the complete quiz and assignment output for the student so that he could immediately concentrate on those problems or concepts which the student had missed on the quiz or assignment.

After some experimentation it was found advantageous to have forty-five minute sessions with about six students at a time. Naturally, students were assigned as much as possible to interview groups according to similarity of demonstrated difficulties.

Near the end of the semester a weekly overall summary of class performance by question was immediately available and invaluable to the instructor. If the quiz and assignment questions were well designed it was easy to spot those problems or concepts that had been generally poorly done or misunderstood. The instructor could then do additional examples or redo some section of the material. This summary was available too late in the project to expect substantial results for all students.

It is our impression that the program also helped to motivate some students. Since all assignments were graded weekly, there was little opportunity to get behind. A composite weekly grade was immediately available so that each student knew his ongoing standard of performance. One complaint several of the better students made was that they felt a couple of computer cards did not adequately represent the amount of time and effort that had gone into completion of the cards.

TABLE I. Comparison of Grade Results.

	4	3	2	1	0	W
X	3.92	12.68	23.68	32.47	22.28	4.96
Y	4.12	13.40	24.74	44.33	9.28	4.12
$\|X-Y\|$	0.20	0.62	1.06	11.86	13.00	0.84

A comparison of the grade results appear in Table I. Grade 0 is the failing grade and W stands for withdrawal from the course early enough in the semester so as not to be assigned a grade. All figures are percentages. X is the average over the four theatre sections which did not use computer-assisted grading. Y is the set of corresponding percentages for the experimental section and the third

row lists the magnitude of the differences in each grade category.

The only substantial differences are in the grade 1 and grade 0 categories. It appears that the program of interviewing those who failed the weekly quizzes resulted in shifting about one half of them into the passing grade.

Since only those who failed each week were interviewed, those in the nonfailing grade categories function as a control both for the effects of the experimental treatment and for the effect of instructional differences. Since there appear to be no instructional effect differences in the other grade categories, we might expect that there were none in the category to which the experimental treatment was applied. Also, a comparison of previous and concurrent grade result patterns for the experimental instructor show no unusual distributions of grades.

Since the experimental theatre section was large and there were no significant differences in the high school mathematics scores between the experimental section and the control sections, the possibility of student differences accounting for the distribution shift is minimal. The final examination was a common three hour written paper which was hand graded for both experimental and control groups.

Our conclusion is that the computer-assisted grading program resulted in substantial improvement in student performance. Since the treatment was on material that the students had demonstrated difficulties with, the expectation is that the treatment would be equally effective on the other grade categories. There are several ways of doing this.

One is to employ the interview procedure with all students who have difficulty with any particular section of the material. This might mean that more interview time would be needed and thus require an additional grader. However, it would be suited for those universities which have a professor give three lectures per week and a graduate student take two small problem sessions.

Another method is to replace the criteria of receiving less than an expected grade. This expected grade could be stated by the student when he enrolls in the course and checked by a correlation with high school performance to see that it is not unrealistic. Then whenever any student drops below his expected grade in a weekly average, he is scheduled an interview appointment to go over those problems he had wrong. Students would be grouped by difficulty rather than by grade category.

We found the project worthwhile and plan to continue it in some of our larger introductory precalculus courses. The problems involved in generating reliable quizzes and assignments for other courses such as the calculus are at this stage too difficult to make the project feasible for smaller campuses where calculus is usually taught to small classes. Such a program to generate materials at these levels of mathematical sophistication would involve considerable experimentation and research.

MISSING INGREDIENTS IN TEACHER TRAINING: ONE REMEDY

MARION I. WALTER AND STEPHEN I. BROWN, Harvard Graduate School of Education

1. Introduction. In discussing the training of mathematics students the need for greater emphasis on nontrivial problem solving at all levels has been stressed. Very few analogous claims have been made, however, for the activity of problem *posing*, and for the interaction between the two. We have created a course which is an invitation for our students in the Master of Arts in Teaching Program to engage in precisely such activity.

The need for such activity became apparent to us when we became aware of the fact that few of our students here and at other colleges have had prior opportunity to:

(1) create mathematics on their own;
(2) learn to pose problems;
(3) make and decide on definitions; and
(4) pursue problems of their own choosing.

In addition, most students have been exposed to only one model of teaching mathematics—the lecture method. Even in cases where the style has not been primarily lecture, the emphasis for the most part has been on understanding and digesting a body of material that usually reflects someone else's organization.

During their four years of learning mathematics in college, most of our students have not had the opportunity to:

(5) discuss the mathematical motivation used to introduce a particular topic in a particular way;
(6) discuss competing motivations for a particular topic;
(7) examine historical development of a mathematical idea;
(8) analyze competing sequencing of topics or courses;
(9) critically examine different solutions (and attempted solutions) to one problem;
(10) critically read and discuss articles in mathematics journals;
(11) investigate the bounds, limitations, and value of a particular structure —beyond merely learning a linear order of axioms, definitions, and theorems;
(12) provide structure for a mass of disorganized conjectures.

While there is no clear separation in categories, suggestions (5)–(12) appear to be means for getting at goals that are implied in (1)–(4).

We are not advocating that all mathematics should be self-initiated nor are we advocating the abolition of the lecture method entirely. We are suggesting that components (1–12) ought, in some form, to be part of a prospective teacher's or any math student's background.

We are not making the absurd claim that every student can do recognized research; however, there is reason to believe that most students can experience the satisfaction of doing something original. There are a host of questions and

problems that nonresearch mathematicians can explore with success, satisfaction, and insight despite the fact that the topics may not be at the forefront of knowledge.

2. Course. Though some of the above experiences could be incorporated in the context of every mathematics course, this in fact is not generally being done. We therefore have attempted to provide our students with *some* of these experiences. We cannot in one semester provide an opportunity for students to engage in *all* these activities in a significant manner. The task is particularly difficult if we appreciate that not only must new attitudes be acquired, but old ones must be relinquished.

As a start, we give the catalogue description of our course below:

Generating and Solving Problems in Mathematics
Half-course (fall term. M., W., 2–4). Marion I. Walter and Stephen I. Brown.

The main purpose of this course is to provide a context which will counteract an approach to mathematics which is characterized by clear organization of content, clearly-posed problems, logical development of definitions, theorems, proofs. We intend instead to provide students with some feeling for mathematics-in-the-making, and will engage in and explore techniques for: generating problems, solving problems, providing structure for a mass of disorganized data, reflecting on the processes used in the above activities, analyzing moments of insight, analyzing "abortive" attempts.

Undergraduate major in mathematics (or its equivalent) and permission of the instructors is required.

The main structural feature of the course that provides a focus for other activities and that seems to get at (either directly or indirectly) many of the goals and activities we previously described is the creation of a class journal. The journal evolves from our students' papers. There are two editorial boards—half the members of the class on one and the rest on the other. Students submit papers to the editorial board of which they are not members. Each board makes written criticism and passes judgment on the papers submitted.

Sources for journal articles include:

(i) problems arising out of class discussions;
(ii) problem suggested by instructors approximately every three weeks—called the problem of the 3 weeks;
(iii) articles on problems appearing in professional journals;
(iv) re-examination of a journal article previously submitted to the editorial board.

The papers could be a student's first attempt at defining, analyzing, or solving a problem. The students could also extend, solve, analyze, criticize one of the topics previously dealt with in the course. We stress that it is not necessary that the problem be solved—though of course some analysis is required. Papers include discussions of false starts, an introspection on insights or misconceptions, and a list of *related* topics and specific problems generated while solving the original problem.

The journal in addition includes a list of interesting problems that come up in class or in small group or editorial board discussions, an abstract for each accepted article, and a list of books or articles either related to specific problems that have been worked on or that provide general background.

There are a variety of interactions both in and out of class. For the purposes of writing for and reviewing the journal, students are encouraged to choose to work alone on some occasions and in small groups cooperatively on others. In addition to working on problems that a student poses himself, each student is required, on at least one occasion during the semester, to work on a problem posed by someone else.

When we teach the entire group in more or less lecture style we have various purposes in mind in addition to transmitting information and providing fertile ground for problem posing. We frequently present a situation, topic, or problem that is open-ended, not well defined, and capable of examination through several points of view. Time is generally allotted during such sessions for individuals, small groups, or the entire class to brainstorm on some ideas presented in the lecture. Some of these explorations eventually emerge as journal articles, and lead to useful, heated, and fertile discussions.

Another activity that students coming to us have had virtually no experience with is reading, discussing, constructively criticizing, or expanding upon articles in professional journals, though they get some of this kind of experience while focussing on the class journal. In order to encourage students to engage in this activity, we have suggested pertinent articles from journals such as the *American Mathematical Monthly*, *The Mathematics Teacher*, *Mathematics Teaching* (British journal) and *Mathematics Magazine*.

3. Selection of mathematics and methods of introducing it. The style and content of the course changes as the term progresses. In the first phase of the course *we* (rather than the students) select a topic which is rich as a potential source for posing and solving problems. After the students have individually and then in small groups "milked" the topic dry (for purposes of problem posing), we begin implicitly to employ some of the problem posing techniques that we have developed. One of these—the "What-if-not" technique—leads to much richer ideas later on in the course when this principle is extended, made explicit, and examined more carefully especially in relationship to other techniques for posing problems. In [1] and [2] we have discussed the What-if-not technique in detail, in one case using a concrete material, and in the other a theorem.

After several weeks, we state explicitly problem posing techniques and introduce various new topics. The topics are intended to encourage students (i) to look at other interesting problems ranging from loosely structured "situations" to some axiomatic considerations, (ii) to find connections among topics that they believe to be unrelated, (iii) to investigate and place in historical perspective a number of basic concepts, (iv) to generate hypotheses and organize

a mass of unorganized data, conjectures, or questions. At this stage the journal is well under way.

During the last phase of the course students examine critically and report on the suggested professional journal articles. They are encouraged to employ techniques acquired over the semester to topics of their own choosing.

Among criteria we have used to select topics are the following:

(i) Students should have *some* machinery available to define and attack problems in the area.

(ii) Topics should lend themselves to examination from a number of different perspectives.

(iii) Although innocent looking on the surface, topics should have unsuspected depth.

(iv) Problems should be such that students can be enticed by easily suggested "situations."

What satisfies these criteria depends upon the background and sophistication of the students. One could select from an endless number of topics or situations that would meet the criteria and satisfy the appetites of students ranging from those in elementary school to those doing doctoral work in mathematics.

We are in the process of analyzing and describing the course in detail and are including the many topics and situations and the manners in which they were employed over the past years. Here we shall instead indicate the flavor and spirit of the course by including six student abstracts. The first three were written by different students on a related topic; the next two by one student who enlarged his previous paper; and the last based on an article in a professional journal.

A Few Results Concerning Primitive Pythagorean Triplets (PPT's): An Alternate Scheme of Classification (by Mr. B_1).

This article develops a method of classifying PPT's according to the value of $c-b$. Formulas, tables, and divisibility results are obtained. The interesting question of which values of $c-b$ are found in PPT's is raised and a partial answer given. For example, where $c-b$ is equal to 3, 4, or 5, the triple a, b, c is *NOT* a *primitive* triple. Enough work is left for the reader to encourage further research by him on this question.

Integral Solutions of $x_1^2+\cdots+x_m^2\leqq n$ (by Mr. T).

This paper combines the results of two previous papers on the described topic. The case for $m=2$ is explained in detail, including some of the attempts that didn't work. The results are then generalized to higher dimensions, and the paper includes some interesting material on formulae for "volumes" of m-dimensional "spheres". Both approximations and an exact recursive formula for the number of solutions are derived. The paper concludes with some notes on brainstorming and problem-solving.

Variations on a Theme by Pythagoras (by Mr. G_1).

This paper looks at three twists on the old theme. Integer-sided 60° triangles, integer points on the ellipse, and the imperfect Pythagorean triples generated by relations of the form

$$a^2+b^2=c^2+k$$

are investigated and some surprising relationships are found.

A Record of Repeated Failures: "Inner Polygons" (by Mr. M_1).

Motivated by the fact that the mid-points of the side of any quadrilateral are vertices of a parallelogram, the author explores the possibility that there might exist regularities among the "inner polygons" (figures formed by repeating the process of connecting the consecutive midpoints of the previous polygons) of a specified n-gon, $n>4$. He offers drawings that seem to support his hypothesis that sides of alternative inner polygons are at least parallel (if the polygons are not indeed similar). His attempted algebraic vertification imposes a restrictive condition on the hypothesis: only certain n-gons have inner polygons which look so nice. Mr. M_1 suggests other aspects for exploration and makes one final conjecture about the inner polygons of n-gons which do not meet his special criterion.

Polygons within Polygons (by Mr. M_1).

Basically an enlargement of his previous article, *A Record of Repeated Failures: "Inner Polygons"*, the present article is an investigation into what happens when mid-points of the sides of a polygon are joined in consecutive order to form a new polygon. A limiting shape for the nth inner polygon is hypothesized, and four transformations which might transform regular polygons into these "limit" polygons are mentioned. An appeal is made for teachers to enable both themselves and their students to become "mathematicians" through exploring interesting problems of their own choosing.

When is the Distributive Law Necessary? (by Mr. M_2).

This paper uses the notion of dual statements in a field, as presented in [3], to develop through a series of propositions a criterion for deciding when the distributive property is necessary for the proof of a field theorem. The paper begins with a discussion of the various difficulties encountered in the course of development of the ideas presented; and a list of brainstorming ideas and unanswered questions is found at the end.

An additional indication of the range of interests is conveyed by the following journal titles:

Some Further Investigations of the Hexed Game Pieces (Mr. B_2)
What If Not "What If Not" (Mrs. H)
An Examination of Gauss' Disquisitions Arithmetical (Miss Mc)
(The student consulted the original French source of 1801)
On the Notion of Mathematical Elegance (Miss C)
An Investigation of the Number of Rational Points Lying on a Circle (Miss W)
Observations on Three Treatments of Continued Fractions (Miss G)
Some Theorems on Polysectioning Geometric Figures (Mr. R)
Toward a General Definition of Similarity in Euclidean 2-Space (Mr. D, Miss R_1, Mr. R)
Determination of Common Integral Solutions to $x^2+y^2\leqq n$ and $x+y=k$ (Mr. B_2, Miss M_1, Miss M_2).

4. Conclusion. We have barely mentioned the problem posing techniques and pedagogical strategy used to employ and encourage further investigation of them. We hope however that our brief description has given the reader some indication of the spirit of the course, the nature of student interaction, and the flexibility of topics that we presented or that interested students.

We are now in the process of evaluating the work of the past three years. We are keeping in mind that our aim was not merely to produce better problem posers and solvers, but also to encourage our students to become more aware

and introspective of their own style of operating both as students and potential teachers. As a first approximation, we feel that this course goes a long way to remedy a number of the defects we have mentioned in the first section.

References

1. Marion I. Walter and Stephen I. Brown, What if not? Math. Teaching, 46 (1969) 38–45.
2. Stephen I. Brown and Marion I. Walter, What if not?: An elaboration and second illustration, Math. Teaching, 51 (1970) 9–17.
3. Stephen I. Brown, Multiplication, addition and duality, Math. Teacher, 59 (1966) 543–50 and 591.

PROBLEMS AND SOLUTIONS

EDITED BY EMORY P. STARKE

All problems (both elementary and advanced) proposed for inclusion in this Department should be sent to E. P. Starke, 1000 Kensington Ave., Plainfield, NJ 07060. Proposers of problems are urged to enclose any solutions or information that will assist the editors. Ordinarily, problems in well-known textbooks and results in generally accessible sources are not appropriate for this Department. No solution (except those accompanying proposals) should be sent to Professor Starke.

ELEMENTARY PROBLEMS

Solutions of Elementary Problems should be sent to Problems Group, Mathematics Department, University of Maine, Orono, ME 04473. To facilitate their consideration, solutions of Elementary Problems in this issue should be typed (with double spacing) and should be mailed before July 31, 1971. Contributors (in the United States) who desire acknowledgment of receipt of their solutions are asked to enclose self-addressed stamped postcards.

E 2263 [1970, 1008]. **Correction.** *Proposed by Bernard McCabe, Bell Comm. Inc., Washington, D.C.*

Suppose an urn contains m balls, each a different color. An observer draws a ball at random, records its color, and replaces it in the urn. He repeats this procedure until some color reappears for the first time. Show that the expected number of drawings is

$$E(m) = \sum_{k=1}^{m} \frac{k(k+1)m!}{(m-k)!\,m^{k+1}},$$

and determine the leading term in the asymptotic expansion.

E 2289. *Proposed by John Corcoran, State University of New York at Buffalo*

Let A be a finite system of independent axioms in a first order logic. Let $\mathcal{E}A$ be the conjunction of the members of A. Then singleton $\mathcal{E}A$ is at least as small as any independent set equivalent to A. Under what conditions is there an independent set equivalent to A and at least as large as any such independent set?

E 2290. *Proposed by E. H. Davis, Kansas State College at Pittsburg*

Describe all polynomials, $p(x, y)$, with real coefficients such that $p(x, y) = p(x+1, y+1)$.

E 2291. *Proposed by Barry Wolk, University of Manitoba*

If $\sum_0 f(n)$ means $f(1)+f(3)+f(5)+\cdots$, show that for all real x

$$\left| \sum_0 n^{-2} \cos(nx) \right| = \sum_0 n^{-2} \cos^2 (nx).$$

E 2292. *Proposed by Stephen Maurer, Phillips Exeter Academy*

Let S be the direct sum $\sum_{i \in I} Z_i$, and P the direct product $\prod_{i \in I} Z_i$, where each Z_i is a copy of the additive group of integers and I is an infinite set. Is the natural image of S in P a direct summand?

E 2293. *Proposed by Erwin Just, Bronx Community College*

Does there exist an infinite set of primes, S, such that whenever $p \in S$ and $q \in S$, we have $(\frac{1}{2}(p-1), \frac{1}{2}(q-1)) = 1$, $(p, q-1) = 1$ and $(p-1, q) = 1$?

E 2294. *Proposed by Douglas Lind, Stanford University*

For which n does the regular n-simplex of side 1 have rational height?

SOLUTIONS OF ELEMENTARY PROBLEMS

Obtuse Triangle Within a Rectangle

E 1150 [1955, 40]. *Proposed by Frank Hawthorne, New York State Education Department*

If three points are selected at random in a rectangle $A \times 2A$, what is the probability that the triangle so determined is obtuse?

Editorial Note. A solution and generalization by Eric Langford may be found in the *Mathematics Magazine* for November 1970, pp. 237–244.

Matrices Whose Elements Are Reciprocals

E 2234 [1970, 403]. *Proposed by P. M. Gibson, University of Alabama at Huntsville*

Prove or disprove: If $A = (a_{ij})$ is a nonsingular $n \times n$ complex matrix, if also

$A^{-1}=(b_{ij})$, and each a_{ij} and b_{ij} is nonzero, then the matrices of reciprocals (a_{ij}^{-1}) and (b_{ij}^{-1}) are singular or nonsingular together. H. Flanders [this MONTHLY, (1966) 270–272] proved this for $n=3$.

Solution by the proposer. The statement is clearly true for $n=1, 2$. We shall show that it is false for all $n>3$. Let $n>3$, and let $A=(a_{ij})=\alpha E+\beta I$, where I is the $n\times n$ identity matrix and E is the $n\times n$ matrix with each entry equal to 1. If $(\alpha+\beta)\alpha\neq 0$, then each a_{ij} is nonzero and

$$(a_{ij}^{-1}) = \frac{1}{\alpha} E - \frac{\beta}{(\alpha+\beta)\alpha} I.$$

It is easy to show that if $\beta\neq 0$ then A is singular if and only if $n\alpha+\beta=0$. This singularity criterion applied to (a_{ij}^{-1}) yields the following:

LEMMA 1. *Let $(\alpha+\beta)\alpha\beta\neq 0$. Then (a_{ij}^{-1}) is singular if and only if $n\alpha+(n-1)\beta=0$.*

If A is nonsingular then

$$A^{-1} = (b_{ij}) = \frac{-\alpha}{(n\alpha+\beta)\beta} E + \frac{1}{\beta} I.$$

If we apply Lemma 1 to A^{-1}, we obtain the following:

LEMMA 2. *Let $((n-1)\alpha+\beta)(n\alpha+\beta)\alpha\beta\neq 0$. Then (b_{ij}^{-1}) is singular if and only if $(n-2)n\alpha+(n-1)\beta=0$.*

In order to show that the statement in the problem is false, it suffices to choose α and β so that the hypothesis of Lemmas 1 and 2 are satisfied while $n\alpha+(n-1)\beta=0$ and $(n-2)n\alpha+(n-1)\beta\neq 0$. A choice that works for all $n>3$ is $\alpha=1-n$, $\beta=n$.

Also solved by D. E. Crabtree, and by E. T. H. Wang.

A Number Theoretic Identity

E 2235 [1970, 522]. *Proposed by Gideon Nettler, Montclair (N. J.) State College*

Prove that

$$\tau^2(d) = \sum_{c|d} \sum_{b|c} \sum_{a|b} \mu^2(a),$$

where $\mu(n)$ is the Möbius function and $\tau(n)$ is the number of divisors of n.

Solution by David Monk, Edinburgh, Scotland. Let

$$f(b) = \sum_{a|b} \mu^2(a), \qquad g(c) = \sum_{b|c} f(b), \qquad h(d) = \sum_{c|d} g(c),$$

so that the relation to be proved is $h(d)=\tau^2(d)$. Since the Möbius function is multiplicative, so are f, g, and h; because the divisor function τ is also multiplicative, it suffices to prove the result for d a prime power. We have, for any prime p and positive integer r,

$$f(p^r) = \mu^2(1) + \mu^2(p) = 1 + 1 = 2; \qquad f(1) = 1,$$

$$g(p^r) = f(1) + \sum_{s=1}^{r} f(p^s) = 1 + 2r; \qquad g(1) = 1,$$

$$h(p^r) = g(1) + \sum_{s=1}^{r} g(p^s) = 1 + \sum_{s=1}^{r} (1 + 2s)$$

$$= 1 + r + r(r + 1) = (r + 1)^2 = \tau^2(p^r); \qquad h(1) = 1.$$

The result follows.

Alternately, and essentially equivalently, we can use known expansions in Dirichlet series: Let

$$F(s) = \sum_{n=1}^{\infty} f(n)n^{-s}, \qquad G(s) = \sum_{n=1}^{\infty} g(n)n^{-s}, \qquad H(s) = \sum_{n=1}^{\infty} h(n)n^{-s}.$$

Then

$$H(s) = \zeta(s)G(s) = \zeta^2(s)F(s) = \zeta^3(s) \cdot \sum_{n=1}^{\infty} \mu^2(n)n^{-s}.$$

Now $\sum_{n=1}^{\infty} \mu^2(n)n^{-s} = \sum_{n=1}^{\infty} |\mu(n)| n^{-s} = \zeta(s)/\zeta(2s)$ [E. C. Titchmarsh, *The theory of the Riemann zeta function*, (Oxford, 1951), eq. (1.2.7)], so

$$H(s) = \zeta^4(s)/\zeta(2s) = \sum_{n=1}^{\infty} \tau^2(n)n^{-s} \qquad [\textit{ibid.}, (1.2.10)],$$

the expansions being valid for $R(s) > 1$. Again the result follows.

Also solved by J. D. Baum, D. M. Cohen, Robert Fray, Ray Glenn, M. G. Greening (Australia), J. Hanumanthachari (India), D. G. Hazelwood, Dean Hickerson, F. W. Humburg, Wells Johnson, Geoffrey Kandall, David Kelly, M. S. Klamkin, L. Kuipers, Douglas Lind, H. Niederreiter, Robert Patenaude, Simeon Reich (Israel), Kenneth Rosen, E. F. Schmeichel, R. Sivaramakrishnan (India), E. A. Smith, Sid Spital, Allen Stenger, E. W. Trost (Switzerland), A. M. Vaidya & A. P. Shah & V. S. Joshi (India), C. S. Venkataraman (India), L. J. Warren, K. M. Wilke, K. L. Yocom, and the proposer.

Editorial Note. In terms of the Dirichlet product, the problem asks us to show that $\tau^2 = I*I*I*\mu^2$, where $I(n) = 1$ for all n. Johnson and Lind show that this is equivalent to the more symmetric $\tau^2*\mu = \mu^2*\tau$, and then prove the latter equality.

Sivaramakrishnan shows that $\tau_k*f = I*I* \cdots *I*f$, where the factor I occurs k times, where f is an arbitrary multiplicative function, and where $\tau_k(n)$ is the number of representations of n as a product of k factors, where order counts. (See M. G. Beumer, *The arithmetic function* $\tau_k(n)$, this MONTHLY 69 (1962), 777–781.) He then shows that $\tau^2 = \tau_3*\mu^2$, so that the problem becomes the special case $k = 3$ and $f = \mu^2$.

Venkataraman proves the following generalization:

$$d^{-2k}\sigma_k^2(d) = \sum_{c|d} \sum_{b|c} \sum_{a|b} \frac{a^k\mu^2(a)}{b^kc^k},$$

where $\sigma_k(n)$ is the sum of the kth powers of the divisors of n. The problem then becomes the special case $k = 0$ of this equality.

The Integral of the Reciprocal of a Polynomial

E 2236 [1970, 522]. *Proposed by M. S. Klamkin, Ford Scientific Laboratory, and D. J. Newman, Yeshiva University*

Show that if the integral of the reciprocal of a nonconstant polynomial is a rational function, then the polynomial must be of the form $(ax+b)^n$.

Solution by G. A. Heuer and C. V. Heuer, Concordia College. If the rational function, in lowest terms, is $f(x)/g(x)$, f and g polynomials, then

$$[g(x)f'(x) - f(x)g'(x)]/[g(x)]^2 = 1/p(x),$$

where p is a polynomial. If $x-r$ is a (possibly complex) factor of the numerator on the left, it divides $[g(x)]^2$. so divides $g(x)$, so divides $f(x)g'(x)$, and therefore $g'(x)$; thus $(x-r)^2 | g(x)$. By induction one finds that if $(x-r)^m$ divides the numerator then $(x-r)^m | g'(x)$ and $(x-r)^{m+1} | g(x)$. Thus the two terms in the numerator separately divide $[g(x)]^2$. Since $f(x)$ and $g(x)$ are relatively prime. $f(x)$ is a constant. Thus $g'(x) | [g(x)]^2$, and every linear factor of $g'(x)$ divides $g(x)$. It follows that if $(x-r)^m | g'(x)$ then $(x-r)^{m+1} | g(x)$. Since the degree of $g(x)$ is only one more than that of $g'(x)$, $g(x)$ cannot have two different linear factors. The desired result follows.

Also solved by Dennis Allen, Jr., Harley Flanders, Michael Goldberg, M. G. Greening (Australia), Harry Lass, Joel Levy, D. J. Peterson, J. R. Ventura, Jr., P. H. Young, and the proposers.

An Extension of a Property of Primitive Roots

E 2238 [1970, 522]. *Proposed by Emanuel Vegh, U. S. Naval Research Laboratory*

Let p be a prime and $t_1, t_2, \cdots, t_n$ a reduced residue system mod $p-1$. If g is a primitive root of p, then it is well known that the integers $g^{t_1}, g^{t_2}, \cdots, g^{t_n}$ are distinct mod p. (In fact, these integers are the primitive roots of p.) Is there an integer h, not a primitive root of p, such that $h^{t_1}, h^{t_2}, \cdots, h^{t_n}$ are distinct mod p?

Solution by David Kelly, Queen's University, Kingston, Ontario. Let P be the assertion that such an h exists for the prime p. Since P is false for $p=2$ we consider only odd primes. Let $R \subset \{1, 2, \cdots, p-2\}$ be the set of reduced residues mod $p-1$. Then P is equivalent to P': "There is a proper divisor d of $p-1$ such that the difference of any two distinct elements of R is not divisible by d." (Then h equals g to the $(p-1)/d$ power where g is a primitive root of p.)

Case (i): $p \equiv 1 \pmod 4$. Let $p=4k+1$. We show that P' is false by showing that for every proper divisor d of $p-1$ there are two distinct elements of R whose difference is divisible by d. Clearly we need only consider the cases where $d=2k$ or $d=4r$ with r a proper divisor of k and k/r odd. In the first case 1 and $2k+1 \in R$, while in the second case 1 and at least one of $4r+1$ or $8r+1 \in R$.

Case (ii): $p \equiv 3 \pmod 4$. Let $p=4k+3$. We show that P' holds with $d=2k+1$. Since all the elements of R are odd, any difference is even and therefore $\neq d$. If $b-a=l(2k+1)$ for $a, b \in R$ and integer $l \geqq 2$, then $b > 4k+2 = p-1$, a contradiction.

In summary, P holds if and only if $p \equiv 3 \pmod 4$.

Also solved by D. M. Bloom, James Bookey, Robert Fray, Bob Gray, M. G. Greening (Australia), C. V. Heuer, F. T. Howard, V. S. Joshi & A. M. Vaidya (India), D. P. Peters, L. J. Warren, and Charles Wexler.

A Relation Existing on Many Finite Sets

E 2239 [1970, 523]. *Proposed by G. Sabbagh, Paris, France*

Let R be a binary relation on a nonvoid set E such that

(1) For no $x \in E$ is xRx true.

(2) For each pair (x, y) of distinct elements of E one and only one of the following relations holds: xRy, yRx.

(3) R is dense, which means: If xRy then there is a $z \in E$ such that xRz and zRy.
Must E be infinite?

I. *Solution by Ralph Seifert, Jr., Hanover College.* If E has just one member and R is the empty relation, then (1)–(3) are satisfied.

If (1)–(3) hold and E has more than one member, then E has at least seven members.

Proof: Suppose there are $x_1, x_2 \in E$ such that x_1Rx_2. Then by (3) there exist x_3 such that $x_1Rx_3Rx_2$, x_4 such that $x_1Rx_4Rx_3$, x_5 such that $x_4Rx_5Rx_3$, x_6 such that $x_3Rx_6Rx_2$, and x_7 such that $x_3Rx_7Rx_6$. Using (1)–(3) it is easy to check that for each $j = 2, 3, \cdots, 7$ we have $x_i \neq x_j$ for all $i = 1, \cdots, (j-1)$. It also follows that each element must relate to at least three elements on its left or right, if it relates on its left or right at all (see x_3).

If E is any finite set with at least seven members, then there is a relation R such that (1)–(3) hold.

Proof: *Case* 1: E has an odd number of members. Suppose $E = \{1, \cdots, p\}$, p an odd number (at least 7). Define R as follows: whenever $1 \leqq n < k \leqq p$, put nRk if $(k-n)$ is even or equal to 1, but different from $(p-1)$; and put kRn if $(k-n)$ is odd or equal to $(p-1)$, but differs from 1. Obviously R satisfies (1) and (2). To prove (3), suppose xRy. First suppose $x < y$, i.e., that $(y-x)$ is even or equals 1. Then if $(y-x) > 2$, we have $xR(x+2)Ry$. If $(y-x) = 2$, we have $xR(x+1)Ry$. If $(y-x) = 1$ and $(x+4) \leqq p$, then $xR(x+4)Ry$. If $(y-x) = 1$ and $x+4 > p$, then $xR(x-(p-4))Ry$ since $p \geqq 7$ and hence $(p-4)$ is odd but not equal to 1. Now suppose $y < x$, i.e., that $(x-y)$ is odd or equals $(p-1)$, but is not 1. If $y \geqq 3$, then $xR(y-2)Ry$. If $x \leqq (p-2)$, then $xR(x+2)Ry$. If $y = 1$ and $x = (p-1)$, then $xRpRy$. If $y = 1$ and $x = p$, then $xR(p-3)Ry$. If $y = 2$ and $x = p$, then $xR1Ry$. This exhausts all possibilities.

Case 2: E has an even number of elements. Suppose $E = \{1, \cdots, p+1\}$, p an odd number. Define R on the set $\{1, \cdots, p\}$ as in Case 1, and also put $nR(p+1)$ whenever $n \leqq p$. Now R satisfies (1) and (2); for (3), suppose xRy. If $x, y \leqq p$, then (since $p \geqq 7$) the previous argument applies; otherwise, $y = p+1$ and $x \leqq p$. But if $x < p$ we have $xR(x+1)R(p+1)$, and if $x = p$ we have $xR1$ $R(p+1)$.

II. *Comment by David Singmaster, Bedford College, England.* In the relation for $p=7$ elements (given in solution I above), call each element a *point*, and each set $L_x=\{y:xRy\}$ a *line*, forming a projective plane of seven points and seven lines. Does each such R define a projective plane? Does each projective plane define such an R?

Also solved by D. N. Adler, P. H. Anderson, A. K. Austin (England), Anders Bager (Denmark), J. D. Baum, L. W. Beineke, Stan Benkoski, J. C. Binz (Switzerland), D. M. Bloom, Walter Bluger, T. C. Brown, Bettye Anne Case, R. O. DeVries, W. F. Fox, Robert Fray & Robert Gilmer, Edward Gade, 3rd, D. P. Geller, G. S. Glazer, Michael Goldberg, Robert Heller, C. V. Heuer & G. A. Heuer, Stephen Hoffman, F. W. Humburg, J. R. Isbell, V. S. Joshi (India), David Kelly, E. E. Kra, W. G. McArthur, E. P. McCravy, David Merriell, Robert Patenaude, D. E. Penney, T. M. Phillips, H. N. Rawal (India), K. B. Reid, David Singmaster (England), N. J. A. Sloane & H. S. Witsenhausen, Richard Stanley, R. C. Steinlage, Ira Sterbakov, Walt Stromquist, Konrad Victor (Israel), Ellis von Eschen, D. R. Woodall (England), P. H. Young, and the proposer.

ADVANCED PROBLEMS

All solutions of Advanced Problems should be sent to J. Barlaz, Rutgers—The State University, New Brunswick, N.J. 08903. Solutions of Advanced Problems in this issue should be typed (with double spacing) on separate, signed sheets and should be mailed before July 31, 1971. Contributors (in the United States) who desire acknowledgment of receipt of their solutions are asked to enclose self-addressed stamped postcards.

An asterisk (*) means neither the proposer nor the editors supplied a solution.

5789*. *Proposed by P. C. Shields, Menlo Park, California*

If A and B are measurable subsets of the unit interval, then $A\times B$ is called a rectangle. Find a measurable subset of the unit square which is not a countable union of rectangles, except for a set of measure zero.

5790.* *Proposed by D. E. Daykin, The University, Reading, England*

Find all nontrivial maps $f:R^2\to R^2$ such that whenever a, b, c are collinear, then $f(a)$, $f(b)$, $f(c)$ are collinear.

5791.* *Proposed by M. Z. Nashed, Georgia Institute of Technology*

For $f:R^3\to R$, $x_0\in R^3$, and nonzero h_1, $h_2\in R^3$, the first and second directional derivatives are defined by

$$\delta f(x_0;\,h_1) = \lim_{t\to 0}\frac{1}{t}\,\{f(x_0+th_1)-f(x_0)\},$$

and

$$\delta^2 f(x_0;\,h_1,\,h_2) = \lim_{t\to 0}\frac{1}{t}\,\{\delta f(x_0+th_2;\,h_1)-\delta f(x_0;\,h_1)\},$$

whenever these limits exist.

Construct a function $f:R^3\to R$ for which $\delta^2 f(x_0;\,h_1,\,h_2)$ is a skew-symmetric

nonzero bilinear form at some $x_0 \in R^3$ (i.e., $\delta^2 f(x_0; h_1, h_2) = -\delta^2 f(x_0; h_2, h_1)$ for all $h_1, h_2 \in R^3$, and $\delta^2 f(x_0; h_1, h_2)$ is linear in h_1 and h_2 separately), or show that such a function does not exist.

5792. *Proposed by W. S. Massey, Yale University*

It is well known that given any finitely presented group G and any integer $n \geqq 4$, there exists a compact, orientable n-manifold M^n such that its fundamental group, $\pi_1(M^n)$, is isomorphic to G. Is an analogous theorem true for non-orientable manifolds? An obvious necessary condition is that G have a subgroup of index 2, since the set of orientation preserving path classes in a non-orientable manifold constitutes a subgroup of its fundamental group which is of index 2.

5793. *Proposed by N. S. Mendelsohn, University of Manitoba*

Let G be a quasigroup with at least two elements and let G satisfy the law $(xy)(y(z(zx))) = y$ for all x, y, z in G. Show that any two distinct elements of G generate a subquasigroup of order 5.

5794. *Proposed by Walter Leighton, University of Missouri*

Consider the (modified) Bessel equation $y'' + p(x)y = 0$, where

$$p(x) = 1 + \frac{1 - 4n^2}{4x^2},$$

and suppose there is a nonnull solution with zeros at $x = a$ and $x = b$ (that is, b is conjugate to a), $0 < a < b$. Prove that if $y(x)$ is any solution such that $y(a) \neq 0$, then the integral

$$\int_a^b [y'^2(x) - p(x)y^2(x)]dx$$

is positive when $n^2 > \frac{1}{4}$ and negative when $n^2 < \frac{1}{4}$.

SOLUTIONS OF ADVANCED PROBLEMS

Kuratowski's 14-Sets

5569 [1968, 199]. *Proposed by Stephen Baron, McGill University*

It is known that at most 14 distinct sets may be obtained from a given subset of a topological space by the operations of closure and complementation applied as often as one likes, in any order. There are many examples of subsets of the real line where this maximum is obtained (e.g., $(0, 1) \cup (1, 2) \cup \{r \mid r \text{ rational}, 2 < r < 3)\} \cup \{4\}$).

Prove or disprove: Any such subset of the reals must contain a set of points which is dense in some interval and whose complement is dense in the same interval. Are there other necessary conditions for such a subset?

Editorial Note. A solution with generalization appears in a paper by E. S. Langford, pp. 362–367, in this issue of the MONTHLY.

5683 [1969, 835]. **Erratum:** On page 86 of the January 1971 issue of this MONTHLY, in line 14 of Solution II, "$f(x_0)=1$" should read "$|f(x_0)|=1$."

A Convergent Series Involving Primes

5724 [1970, 313]. *Proposed by Paul Erdös, University College of Swansea, Wales*

Denote by $P(n)$ the largest prime factor of n. Let $f(n)$ be an increasing function of n so that $\sum_{n=1}^{\infty} 1/nf(n)<\infty$. Prove that $\sum_{n=1}^{\infty} 1/nf(P(n))<\infty$.

Solution by P. T. Bateman, University of Illinois. For positive integral m let us put

$$G(m)=\sum_{P(n)\leq m}\frac{1}{n}=\prod_{p\leq m}\left(1+\frac{1}{p}+\frac{1}{p^2}+\cdots\right).$$

By an elementary theorem on the distribution of primes,

$$G(m)=\prod_{p\leq m}\left(1-\frac{1}{p}\right)^{-1}\leq K\log(m+1),$$

where K is an absolute constant. Hence

$$\begin{aligned}\sum_{P(n)\leq N}\frac{1}{nf(P(n))}&=\sum_{m=1}^{N}\frac{G(m)-G(m-1)}{f(m)}\\&=\sum_{m=1}^{N-1}G(m)\left\{\frac{1}{f(m)}-\frac{1}{f(m+1)}\right\}+\frac{G(N)}{f(N)}\\&\leq\sum_{m=1}^{N-1}K\log(m+1)\left\{\frac{1}{f(m)}-\frac{1}{f(m+1)}\right\}+\frac{K\log(N+1)}{f(N)}\\&=\sum_{m=1}^{N}\frac{K\log(m+1)-K\log m}{f(m)}<\sum_{m=1}^{N}\frac{K}{mf(m)}.\end{aligned}$$

Hence the desired result follows.

Also solved by J. H. E. Cohn (England), J. H. van Lint (Netherlands), and Dieter Wolke (Germany).

Partitioning Semigroups

5726 [1970, 313]. *Proposed by D. P. Allen, Jr., Bell Telephone Laboratories, Holmdel, N. J.*

Let X be a finite nonempty set, let $\sum X[\Delta X]$ denote the free [free commutative] semigroup on X, and let $\phi:\sum X\to\Delta X$ denote the natural homomorphism. Do there exist semigroups T and T' of $\sum X$ with $T\cup T'=\sum X$ and $T\cap T'=\varnothing$ such that (i) $\phi(T)\cap\phi(T')\neq\varnothing$ and (ii) neither T nor T' contains a right or a left ideal of $\sum X$?

Solution by Carl Eberhart and Wiley Williams, University of Louisville. We do not restrict ourselves to considering X finite. The answer is yes if $|X|>1$ (if $|X|=1$, then $\sum X=\Delta X$, so condition (i) contradicts the conditions on T and T'.) Suppose $|X|>1$ and let a and b be distinct elements of X. Let $T=\{W\in\sum X \mid W$ contains more a's than b's$\}\cup\{(ab)^n\}_{n=1}^{\infty}$ and let $T'=S\setminus T$ $=\{W\in\sum X \mid W$ contains no more a's than b's and $W\neq(ab)^n$ for any $n\}$. Easily, T is a subsemigroup of $\sum X$. To show that T' is a subsemigroup, suppose $WW'\in T'$. Then WW' contains no more a's than b's. If $WW'=(ab)^n$ for some n, then W begins with a and hence must end with b. So $W=(ab)^m$ for some m, a contradiction to $W\in T'$. So $WW'\neq(ab)^n$ for any n, and hence $WW'\in T'$.

Clearly T and T' satisfy $T\cup T'=\sum X$ and $T\cap T'=\varnothing$. Since $\phi(ab)=\phi(ba)$ they also satisfy (i). To show (ii), suppose R is a right ideal contained in T. If $q\in R$, then $q\sum X\subset R\subset T$. However, if m is the number of a's in q and n is the number of b's in q, then $qb^{n+1-m}\in T'$, and hence $R\not\subset T$. Similar arguments show that neither T nor T' can contain a right or a left ideal of $\sum X$.

Also solved by A. A. Jagers (Netherlands), Stephen Meskin, and Roy Olson.

Invertibility of Formal Matrix Determinants

5727 [1970, 409]. *Proposed by P. R. Halmos, University of Hawaii*

A matrix

$$M=\begin{vmatrix} A & B \\ C & D \end{vmatrix}$$

whose entries are matrices has four formal determinants: $AD-BC$, $AD-CB$, $DA-BC$, $DA-CB$. (In this sense a matrix of size n, instead of 2, has $(n!)^{n!}$ formal determinants.) Prove or disprove the following two assertions. (1) If M is invertible, then at least one of its formal determinants is invertible. (2) If all the formal determinants of M are invertible, then so is M.

Solution by Robert Kimble, Jr., United States Naval Academy. Both assertions are false. Consider

$$(1)\qquad M=\begin{vmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{vmatrix} \quad \text{and (2)} \quad M=\begin{vmatrix} 1 & 0 & 2 & 0 \\ 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & 1 \\ 0 & 2 & 0 & 1 \end{vmatrix}.$$

In (1)

$$A=\begin{vmatrix} 1 & 0 \\ 0 & 0 \end{vmatrix},\quad B=\begin{vmatrix} 0 & 0 \\ 1 & 0 \end{vmatrix},\quad C=\begin{vmatrix} 0 & 1 \\ 0 & 0 \end{vmatrix},\quad \text{and}\quad D=\begin{vmatrix} 0 & 0 \\ 0 & 1 \end{vmatrix}.$$

$$AD-BC=DA-BC=\begin{vmatrix} 0 & 0 \\ 0 & -1 \end{vmatrix} \quad \text{and} \quad AD-CB=DA-CB=\begin{vmatrix} -1 & 0 \\ 0 & 0 \end{vmatrix}$$

and $M^2=I$; M is invertible while all of the four formal determinants are not invertible.

In (2)

$$A=\begin{vmatrix}1 & 0\\ 1 & 0\end{vmatrix},\quad B=\begin{vmatrix}2 & 0\\ 0 & 1\end{vmatrix},\quad C=\begin{vmatrix}1 & 0\\ 0 & 2\end{vmatrix},\quad\text{and}\quad D=\begin{vmatrix}0 & 1\\ 0 & 1\end{vmatrix};$$

$$AD-BC=AD-CB=\begin{vmatrix}-2 & 1\\ 0 & -1\end{vmatrix}\quad\text{and}\quad DA-BC=DA-CB=\begin{vmatrix}-1 & 0\\ 1 & -2\end{vmatrix}.$$

Thus all four formal determinants are invertible while M is not.

Also solved by Einar Andresen (Norway), Marcia Ascher, Alfred Gray, G. F. Votruba, L. C. Washington, and the proposer.

Powers notes that with additional hypotheses on the submatrices both assertions are true. See R. F. Gantmacher, *The Theory of Matrices*, V. I (Chelsea, New York, 1959) pp. 45–46.

Partitioning Sets in Non-Parallel Hyperplanes

5728 [1970, 410]. *Proposed by F. N. Fritsch, University of California, Livermore*

Let $X=\{\mathrm{x}_1, \cdots, \mathrm{x}_N\}$ be a finite subset of E_n with $N\leqq 2n$. Show that there exist two hyperplanes $\mathcal{H}$ and $\mathcal{K}$ such that: (i) $X\subset\mathcal{H}\cup\mathcal{K}$; (ii) $\mathcal{H}$ and $\mathcal{K}$ are not parallel.

Solution by G. A. Heuer and C. V. Heuer, Concordia College. Let H_1 be a hyperplane containing $x_1, \cdots, x_n$, and H_2 one containing the points of X not in H_1. If $X\subset H_1$, we may choose H_2 not parallel to H_1. If H_2 is parallel to H_1, choose n points of X including some from each of H_1 and H_2. Let H_3 be a hyperplane containing these n points, and H_4 one containing those not in H_3. Note that H_3 is not parallel to H_1. If H_4 and H_3 are parallel, consider the four $(n-2)$-dimensional intersections

$$D_{13}=H_1\cap H_3,\quad D_{14}=H_1\cap H_4,\quad D_{23}=H_2\cap H_3,\quad D_{24}=H_2\cap H_4,$$

whose union contains X. We now show there are nonparallel hyperplanes H_{1423} and H_{1324} containing $D_{14}\cup D_{23}$ and $D_{13}\cup D_{24}$, respectively. Let the equations be: $H_1{:}\,a_{11}x_1+\cdots+a_{1n}x_n=b_1$ (or in vector notation, $a_1\cdot\mathrm{x}=b_1$); $H_2{:}\,a_1\cdot\mathrm{x}=b_2$; $H_3{:}\,a_3\cdot\mathrm{x}=b_3$; $H_4{:}\,a_3\cdot\mathrm{x}=b_4$. Note that $b_2\neq b_1$ and $b_4\neq b_3$. For any k, the hyperplane with equation $(a_1+ka_3)\cdot\mathrm{x}=b_1+kb_4$ contains D_{14}, and the one with equation $(a_1+ka_3)\cdot\mathrm{x}=b_2+kb_3$ contains D_{23}. With $k_{1423}=(b_1-b_2)/(b_3-b_4)$ these coincide; hence the desired H_{1423}. Similarly one finds H_{1324} with equation $(a_1+k_{1324}a_3)\cdot\mathrm{x}=b_1+k_{1324}b_3$, where $k_{1324}=-k_{1423}$ ($\neq 0$ since $b_1\neq b_2$). Since a_1 and a_3 are linearly independent, so are $a_1+k_{1423}a_3$ and $a_1-k_{1423}a_3$, and thus H_{1423} and H_{1324} are not parallel.

Also solved by D. Ž. Djoković, David Kelly, I. P. Mysovskikh (U.S.S.R.), S. K. Thomason, and the proposer.

Homeomorphism of the Real Line with Upper Limit Topology

5730 [1970, 410]. *Proposed by K. D. Juhlin, University of Illinois*

Let E_1^u be the real line with the upper limit topology (basis consists of all sets of the form $(a, b]$). Is $(0, 1]$, with the topology induced by the topology on E_1^u homeomorphic to E_1^u (i.e., is E_1^u homeomorphic to one of its basic open sets?)?

Solution by Mark Yu, Columbia University. E_1^u and $(0, 1]$ are homeomorphic. As a matter of fact, they are both homeomorphic to a countably infinite union of pairwise disjoint basic open sets, viz.,

$$E_i^u = \bigcup_{i=-\infty}^{+\infty} (i, i+1], \quad (0, 1] = \bigcup_{i=1}^{\infty} \left(\frac{1}{2^i}, \frac{1}{2^{i-1}}\right].$$

By pairing off the basic open sets in the above two decompositions, one may construct a naturally bijective function between E_1^u and $(0, 1]$ which is continuous in both directions.

Also solved by D. F. Behan, Helen F. Cullen, T. E. Gantner, Bob Gray, D. A. Hejhal, G. A. Heuer, A. A. Jagers (Netherlands), Roger Marty for Cleveland State University Problem Solving Group, P. R. Meyer, T. M. Phillips, Peter Ross, G. C. Schmidt, G. P. Speck, T. A. Straeter, and the proposer.

Representation of Elements in Product Spaces

5731 [1970, 410]. *Proposed by L. N. Childs, State University of New York at Albany*

(a) Let M, N be finite dimensional vector spaces over a field R. Let S be a field extension of R and let M_s, N_s denote the S-spaces generated by M and N ($M_s = M \otimes_R S$, etc.). Prove the following result:

(*) If x in $M \otimes_R N$ has the property that for some S as above, $x \otimes 1 = x_s$ in $M_s \otimes_R N_s$ has the form $x_s = y \otimes_S z$, y in M_s, z in N_s, then $x = u \otimes v$, u in M, v in N.

(b) Assume now only that M, N are finitely generated free modules over a commutative ring with unity R, and let S be an R-algebra which is a finitely generated projective R-module. Find conditions on R so that (*) is still true.

Solution by the proposer. Part (a) is a direct application of the fact that the general element of $M \otimes_R N$ is a sum. We consider part (b).

For R any commutative ring with identity, let M, N be free R-modules with bases e_i, $i = 1, \cdots, m$; f_j, $j = 1, \cdots, n$, respectively. Then $x = \sum r_{ij}(e_i \otimes f_j)$ in $M \otimes_R N$ has the form $x = u \otimes v$, u in M, v in N, if and only if $x = \sum_i s_i e_i \otimes \sum_j t_j f_j = \sum_{i,j} s_i t_j (e_i \otimes f_j)$ for some s_i, t_j in R, if and only if the $m \times n$ matrices (r_{ij}) and $(s_i t_j)$ are equal, if and only if the row space V of (r_{ij}) is a free submodule of rank 1 of R^n (the free R-module of rank n consisting of n-tuples). Suppose S is an R-algebra which is a finitely generated projective R-module (more generally one could assume S is a faithfully flat R-algebra) and $x_s = y \otimes z$, y in M_s, z in N_s. Then the S-subspace V_s of S^n spanned by V is a free S-module of rank 1, hence V is a rank one projective R-module. Thus, if one assumes of R that every rank

one projective R-module is free, then V is a free R-module of rank one and (*) holds.

Without the assumption that rank one projective R-modules are free, we have the following example (this solver doesn't know whether (*) holds if *and only if* rank one R-projectives are free): Let $R=\mathbf{Z}(\sqrt{-5})$ be the familiar example of a number ring without unique factorization. Let $M=N=Re_1\otimes Re_2$ be a free module of rank 2 with basis e_1 and e_2. Let x in $M\otimes_R N$ correspond as in the first paragraph of this solution to the matrix

$$(r_{ij}) = \begin{pmatrix} 2+\sqrt{-5} & 3 \\ 3 & 2-\sqrt{-5} \end{pmatrix}.$$

Then the row space of V of (r_{ij}) is not a free R-module of rank 1 because the ideals $I=(2+\sqrt{-5})R+3R$ and $J=(2-\sqrt{-5})R+3R$ are not principal. But for S a Dedekind overring of R in which IS and JS are principal ideals (e.g., let S be the ring of integers of $Q(\sqrt{-5})(\sqrt{-1})$) it may be verified that V_s, the S-space spanned by V is a free S-module of rank one. Thus $x_s=y\otimes z$ in $M_s\otimes_S N_s$ but x in $M\otimes_R N$ is not of that form.

For R a field, the solution to (*) specializes to noting that any $m\times n$ matrix of rank one is the product of a column matrix and a row matrix.

REVIEWS

EDITED BY J. ARTHUR SEEBACH, JR. AND LYNN A. STEEN
with the assistance of the mathematics departments of St. Olaf and Carleton Colleges

COLLABORATING EDITOR FOR FILMS: SEYMOUR SCHUSTER, Carleton College

Printed materials for review should be sent to: Book Review Editor, American Mathematical Monthly, St. Olaf College, Northfield, MN 55057. Films and correspondence relating to films should be sent to Seymour Schuster, Carleton College, Northfield, MN 55057.

*All unsigned material is written by the editors. A boldface capital **C** in the margin indicates that a review is based in part on classroom use. Professors willing to write such a review should inform the editor in order to avoid duplication.*

Lebesgue's Theory of Integration. Its Origins and Development. By Thomas Hawkins. University of Wisconsin Press, Madison, 1970. 242 pp. $12.50. (Telegraphic Review, December 1970.)

Almost all books on the history of mathematics are concerned with the development of the subject prior to the nineteenth century. The volume under review presents by contrast the ideas which during the nineteenth century led to the definition of the Lebesgue integral. The book starts with the mention of the problem of the vibrating string in applied mathematics and the fact that the solution of the differential equation presented, among other things, the question

of what is an arbitrary function and matters of convergence of Fourier series. Following highlights are: Cauchy's definition of an integral of a function as a limit of sum, which he shows converges for continuous functions; Dirichlet's contributions to the theory of Fourier series which led him to the definition of a function, not via analytic expressions, but as the correspondence between the elements of one set (points of an interval) with another set (real numbers); the Riemann definition of a definite integral which uses the limiting process of Cauchy, which in turn defines a class of functions for which the process converges; the attempt to connect topological and measure properties of the set of points of discontinuity of an integrable function leading Cantor to the formulation of his set theory; the Jordan theory of measure based on coverings by finite sets of intervals; the Borel theory of measure recognizing the importance of complete additivity of set functions which led Lebesgue to the use of infinite coverings instead of finite ones for defining measurability of sets, to a definition of measurable functions, and to a definition of integral based on subdivisions of the range of the function instead of the domain as in the Riemann integral and to the proof of the fact that the process leads to an integral for all bounded measurable functions as well as an easy extension to unbounded functions. There are digressions, for instance the derivability of continuous functions and functions of bounded variation, the integral as the inverse of the derivative, the Fubini theorem on iterated integrals connected with the definition of an integral as the measure of a plane set of points, and so on. The application of Lebesgue integrals is mainly to the theory of orthogonal systems of functions, a generalization of Fourier series. There is a brief chapter on the Riemann-Stieltjes (RS) and a report on the Radon development of the Lebesgue-Stieltjes (LS) integral. It is interesting that Lebesgue showed how to reduce in a rather complicated way an RS-integral to a Lebesgue integral, but apparently missed the fact that his definition of integration of a bounded measurable function is the simple RS-integral $\int_m^M y d\mu(y)$, where $m<f<M$ and $\mu(y) = \text{meas } E\ (f \leqq y)$.

This is an interesting book covering the development of an important part of real analysis. It is valuable to the worker in the field, as it brings out a number of ideas and results which have been more or less forgotten. It can be recommended highly to students who are getting their introduction to Lebesgue integration, particularly because it shows how an important mathematical idea developes, sometimes slowly, until it becomes an esthetically satisfying structure.

T. H. Hildebrandt, University of Michigan

Linear Functional Analysis: Introduction to Lebesgue Integration and Infinite-Dimensional Problems. By Bernard Epstein. Saunders, Philadelphia, 1970. 229 pp. $9.50. (Telegraphic Review, March 1970.)

In the author's words, this book is designed to "(introduce) the reader to a very limited, and comparatively elementary, portion of the vast field of functional analysis in such a manner that he will see that the development of the

subject matter was stimulated by significant problems of concrete, or 'hard,' analysis and that, conversely, the abstract theory may often be effectively employed in the study of specific concrete problems." The book appears to this reviewer to be suitable for use as a textbook by graduate students, either concurrently with, or even before the usual intensive real variables course. Strong undergraduates may find the book useful either for independent study or in a topics course.

The author has tried, successfully, to relate concepts to the geometric or analytic situations out of which they arose. He proceeds from the concrete to the abstract. His explanations are unusually clear and complete. The text and exercises incorporate a large number of examples, which range from the *tournants dangereux* variety to some which indicate the scope and applicability of the text.

Professor Epstein's book includes chapters on metric spaces, Lebesgue measure and integration, normed linear spaces, linear functionals, operators, operators on finite-dimensional spaces, and spectral theory. If there is a single goal which the author has set for himself, it is the exposition of the Fredholm Theory in a context in which both the abstract theory and its interpretation in terms of integral equations are intelligible. Often the author restricts his formal treatment to important special cases of general results (e.g., spectral theory for compact operators, the Fredholm alternative for hermitian compact operators, Lebesgue measure on the line); he then states more general results. This technique keeps the book a manageable length, an "introduction" to functional analysis. Actually, the author covers more ground than he claims.

Epstein's book should appeal to a wide class of readers, both as a means of acquainting themselves with basic ideas of functional analysis and as preparation for such treatises as Riesz-Nagy's "Functional Analysis."

A. J. SILBERGER, Bowdoin College

Introductory Real Analysis. By A. N. Kolmogorov and S. V. Fomin. Revised English edition translated and edited by R. A. Silverman. Prentice-Hall, Englewood Cliffs, N. J., 1970. 415 pp. $13.95. (Telegraphic Review, November 1970.)

The translator describes this revision of the Russian original as "a comprehensive, but manageably proportioned and entirely elementary introduction to real and functional analysis, from a consistently modern point of view." This is a reasonable claim for students well grounded in rigorous advanced calculus. In fact, with its readable, carefully developed, and well-motivated presentation, its wealth of provocative and contemporary concepts, and its judiciously selected and arranged problems, this book should be a delight to the teacher and a realistic challenge to the student.

The standard basic material is all here, including set fundamentals, topological and linear spaces, linear functions and operators, and measure, integration, and differentiation. The style of exposition is illustrated by the treatment

of measure. This is started with outer and inner measure for subsets of the unit square in the Euclidean plane, and then extended to unbounded sets. The abstraction to general measure begins with semirings with unit, continues to semirings without unit, extends to generated rings, and finally includes the abstract generalization of Lebesgue measure in the plane. The Lebesgue integral is presented by means of uniformly convergent sequences of simple functions and a passage to a limit.

Some idea of the extent of the coverage achieved in 389 pages of text can be gotten from two brief selected lists of topics. A few "standard" items are Arzelà's theorem, the Hahn-Banach theorem, weak and weak-star topologies, the spectrum of an operator, the Lebesgue dominated convergence theorem, the Radon-Nikodym theorem, Fubini's theorem, and the Riesz representation theorem. Some of the "less standard" topics treated are topological linear spaces, generalized functions, and Helly's convergence theorem. The L^p spaces are studied only for $p=1$ and 2.

The problems (some 350) are well designed to illustrate the textual material through specific examples and to present such additional concepts as the Hölder and Minkowski inequalities, partial ordering of topologies and norms, Hamel bases, the closed graph theorem, spectral radius, and Jordan measure. There appear to be remarkably few typographical errors, and the index is unusually complete and useful. This book deserves an enthusiastic welcome and wide use.

J. M. H. OLMSTED, Southern Illinois University at Carbondale

Mathematical Geodesy. By Martin Hotine. ESSA Monograph 2. United States Government Printing Office, Washington, 1969. xvi+416 pp. $5.50.

This beautiful book will open the eyes of any geometer who thinks that "mapping" is properly represented by the short discussion of "cartography" contained in the usual differential geometry texts. Although Gauss was directly inspired to found the modern theory of surfaces by the practical problems of surveying Hanover, the succeeding 150 years have seen geodesy disappear from the field of interest of most differential geometers. After examining this fascinating book, we are at a loss to explain why.

To Gauss, geodesy was two-dimensional. Although surveying instruments operated in three dimensions, all lines of sight were projected onto an idealized surface and elevations were treated in a way quite different from horizontal measurements. Yet, within a few decades after Gauss' papers, surveying parties were standing at Dante's View in Death Valley from which they could see both Badwater and the top of Mt. Whitney, points differing in elevation by almost 15,000 feet. And now, with artificial satellites in common use as a geodetic tool, the vertical measurement is often of the same order of magnitude as or significantly greater than the horizontal measurement. Geodesy is now unavoidably a three-dimensional science, studying the Riemannian three-manifold-with-boundary formed by that part of the material universe on the Earth's surface and above it. Various instrumental techniques give rise to different local coordi-

nate systems in this three-manifold and it is necessary to compare them (e.g., astronomical vs. ellipsoidal vertical at the Earth's surface). Thus, a great part of the book is concerned with describing standardized local coordinate systems and to detailing coordinate transformations amongst them. Although calculations for specific "real problems" are never introduced, it is clear that such problems were always in Hotine's mind (indeed, this book was written at the close of a very distinguished career combining both "theoretical" and "applied" geodesy).

The first 66 pages are a condensed introduction to the differential geometry of Euclidean three-space. It is no exaggeration to state that all of the classical theory of curves and surfaces finds application in geodesy: covariant differentiation, conformal transformation, Mainardi-Codazzi equations, and the Gauss-Bonnet theorem are described, as well as just about anything else from classical differential geometry that one could want. By "classical" we mean in the nineteenth-century spirit (and by no means pejoratively). In this spirit, the exposition is made entirely in the tensor calculus. Every geometer knows that when it is time to work on a "dirty-hands" problem, tensor calculus is frequently the method of choice. Yet we wonder if the use of twentieth-century techniques of differential geometry, especially É. Cartan's method of moving frames, could have the same salutory effect on physical mathematics that tensor analysis did when it replaced cartesian methods. Cartan's own writings on the theory of surfaces in three-space give convincing demonstrations of the practicality and elegant transparency of Cartan's methods. Cartan himself warned us long ago of *"les débauches d'indices masquent une réalité géométrique souvent très simple."* We know that the introductory material on differential geometry could have been simplified and drastically shortened by use of the method of moving frames and the structural equations of Cartan. We feel certain that similar simplifications and a closer rapport with the physical situation underlying the geometry would have resulted if Cartan's methods had been employed throughout the book. Hotine was aware of these methods, for in the Preface he recommends H. Guggenheimer's *Differential Geometry*, written in the Cartan spirit. To a reader unfamiliar with Cartan's approach to differential geometry, we also recommend Guggenheimer's book or, better yet, H. Flanders' *Differential Forms* which offers a simple, straightforward exposition and also *applications* to physical problems ordinarily treated by tensor calculus.

For selfish reasons, we hope that *Mathematical Geodesy* and other works like it will be "translated" into the Cartan calculus. Physical scientists will not be the only ones to profit thereby. We think that geometers still have further inspiration to draw from geodesy. Two millenia of response to cartographic demands culminated for mathematics in the *Theorema Egregium* of Gauss. Surely the variety of modern geodesy has much more left to give to modern geometry than just the words *map*, *chart*, and *atlas*. We strongly suspect that the magnificent nugget discovered by Gauss is only one of many hidden in a rich vein.

We are delighted to have *Mathematical Geodesy* to use. It is a book for all

mathematicians. It belongs in every college library. No longer can any self-respecting differential geometry textbook be written without reference to it. And in times of pretentious prices for shoddily produced books, this book, immaculately printed on high-quality paper and well-bound, costs just \$5.50 (only $1\frac{1}{4}$¢ per page). Let the reader draw his own inferences as to the benefits of a government scientific printing establishment.

NATHANIEL GROSSMAN, University of California, Los Angeles

C *Linear Algebra*. By Robert R. Stoll and Edward T. Wong. Academic Press, New York, 1968. x+326 pp. \$8.50.

A strong, basic text on linear algebra, the book begins with the usual axioms for general vector spaces and develops the basic concepts in the first 5 chapters. The heart of the book is in Chapters 6–8, which provides a well-motivated presentation of the characteristic minimum polynomials, and related theory, leading to a lucidly written section on extremal properties of characteristic values, and climaxing in the spectral decomposition theorem (p. 266). Finally, the results of Chapters 6–8 are put to effective use in solving certain problems in applied mathematics, a unique feature of the book, written with the help of carefully chosen experts in the fields of economics, chemistry, and physics. For the most part the exercises are reasonable, interesting, and excellently chosen; computation-type problems are included in practically every section.

The book was written for the more mature undergraduate, in that abstract reasoning is used as the principal mode of communication; it is thus unsuitable for general use at the pre-advanced-calculus level (the authors say only that the book was intended for undergraduates). Regarding length, in classes at the University of Oklahoma most of us who have taught a semester course from this book are successful in covering only the essential topics from Chapters 1–5 and a sampling from Chapter 6.

Although this reviewer concurs wholeheartedly with the choice of topics, the style in the early chapters is such that it requires meticulous study by even the mature student (for example, one finds a general proof of the existence of a basis—involving Zorn's lemma—as only the third topic covered). Further, certain ideas that should be eminently transparent seem to be obscured either by over-information or location with "optional" material (case in point: the solvability of an $n\times n$ system closes the discussion on annihilators). Conspicuously omitted is the very useful theorem concerning the existence of an orthogonal transformation which maps a given orthonormal basis $\{\alpha_1, \cdots, \alpha_n\}$ onto another $\{\beta_1, \cdots, \beta_n\}$ (the formula for the desired transformation is simply $\alpha A = \sum_{i=1}^n (\alpha | \alpha_i)\beta_i$; this result should replace the useless Theorem 7.2).

A more serious fault of the early chapters, in view of the subject matter, is the ponderous treatment of matrices and determinants. The authors actually seem to belittle that area in the remark on p. 145: "For some, matrices have a life of their own, that is, an existence apart from representing linear transformations." Insisting on expressions such as $(a_{ij})(b_{ij})$ to represent matrix multiplica-

tion, the authors choose not to use standard notation, and matrix proofs are avoided throughout. Determinants are developed axiomatically, with an unduly involved proof used for the property det $AB = \det A \det B$ (depending on the uniqueness of the determinant function as opposed to the use of elementary matrices to reduce the problem to diagonal matrices). [At this point, the reviewer would like to point out an alternate development of determinants for courses in linear algebra, which is as rigorous as, but more teachable than, the axiomatic formulation: Let $D(A) = C(A)^t$, where $C(A)$ is the matrix consisting of the cofactors of the elements of A (t denotes "transpose"); the property $A \cdot D(A) = D(A) \cdot A = aI$ (I = identity) and the definition det $A = a$ can be easily developed inductively and illustrated, motivating the student from the beginning to master the Laplace expansion and streamlining many proofs (most relying on simple induction).] It is the reviewer's opinion that, intentionally or otherwise, the authors place themselves in the category of those who, along with hosts of other writers on linear algebra, are guilty of "killing determinants" and, depending on one's point of view, this will undoubtedly serve to identify them as either heroes or villains.

D. C. KAY, University of Oklahoma

C *A First Course in Linear Algebra*. By Daniel Zelinsky. Academic Press, New York, 1968. viii+266 pp. $6.50.

This is an excellent text for a one-term linear algebra course to follow the calculus of one variable and to precede the calculus of several variables. For the average freshman it is neither too computational (as many chapters inserted in calculus books are) nor too abstract (as many linear algebra texts are).

This text begins with ordered triples which are introduced both geometrically and algebraically. Gradually new ideas and generalizations are presented, usually motivated by previous concepts and concrete illustrations. For example, though the concept of basis is not introduced until page 145, many systems of homogeneous equations had previously been solved and their solutions expressed as linear combinations of certain specific solutions. Ideas are not introduced and then forgotten, but rather are used repeatedly throughout the rest of the book. Many theorems and problems bring together several ideas, as in the theorem giving various conditions for the invertibility of a matrix. The exercises, though not original, are well chosen to illustrate the concepts and theoretical results. After the student has met such words as linear function, linear combination, kernel, image, rank, and n-tuple, and has become familiar with the use of definitions, theorems, and proofs, he is prepared for the first really abstract chapter of the book, the chapter on dimension.

The choice of material to be included and the concepts to be omitted in a one-term course aimed at freshman is always a problem. This reviewer regrets (1) the omission of the concepts of homomorphism and isomorphism; (2) that the concept of vector subspace is not related to the concept of vector space (the latter is an optional section); (3) the small amount of material on eigenvectors;

(4) that the concept of basis is introduced before linear independence; and (5) the absence of any practical applications. In addition, the use of four simultaneous numbering systems in each section, although explained, is still confusing. The last two chapters are considerably less well-done than the others.

However, the generally well-organized, readable style accomplishes its aims for the beginning student so well that the shortcomings listed above seem minor. Gradually, but relentlessly, the student rather willingly increases his mathematical intuition and understanding.

Proof of the suitability of the book for its intended purpose is the fact that for three successive years in our freshman mathematics sequence we have followed three different beginning calculus books with this text for linear algebra. A more successful book for its intended audience would be hard to find.

GEORGE NIELSEN AND JEAN CALLOWAY, Kalamazoo College

Elementary Cryptanalysis—A Mathematical Approach. By Abraham Sinkov. New Mathematical Library 22. Published for the monograph project of the School Mathematics Study Group. Random House/Singer, New York, 1968. 198 pp. $1.95 (paper). (Telegraphic Review, November 1969.)

The concept of this book is to use the study of cryptanalysis as a vehicle to introduce mathematical concepts, and a happy concept it is. Everyone knows how fascinating secret messages are, and how like magic it is to read them despite their ciphering. Sinkov's attempts to unravel ciphers call upon probability, modular arithmetic, matrix multiplication, permutations and statistics. Each of these is introduced *de novo*, together with such related ideas as are needed, in a way that an industrious high school student can handle. Included are 87 challenging problems.

The statistical test used is related to the chi-squared. Tables are not used to interpret the statistic; a loose appeal to intuition is used instead. I regard it as unfortunate that a factor was not introduced to bring the I. C. (index of coincidence) into a range where intuition can work better. The proper factor is the size of the alphabet, 26 or 625 or 676 in the three different applications of Sinkov. Then in every case the expected value at random is 1.0 rather than variously .038 or .0016 or .0015 as in the text. But in making this choice the author was within his rights.

The solutions of the cryptanalytic problems in this book are advanced for their mathematical interest. There are better solutions in most cases based on properties of language. This is nowhere alluded to, but is of some interest on its own. It leaves an opportunity for an ingenious explorer to find new techniques on his own, which would then open new and deep questions about which techniques are better, which will work on the least data, and how to combine the techniques on refractory problems.

H. H. CAMPAIGNE, Slippery Rock, Pa.

TELEGRAPHIC REVIEWS

Telegraphic reviews are intended to give prompt notice of books, with sufficient information to assist in deciding whether to order an examination copy or suggest library purchase. Possible uses are indicated as follows:

B = college bookstore stock   L = library purchase
P = professional reading   S = supplementary reading
T = textbook   E = teacher education
13 to 18 = freshman to second year graduate level usage
1 to 4 = approximate time in semesters to cover text
* = positive emphasis   ? = negative emphasis

Books on high school material (pre-calculus) are denoted REMEDIAL, and normally receive telegraphic reviews only if they are written for college students. Publishers are denoted by the standard abbreviations used in *Books in Print*, which gives complete addresses.

ALGEBRA, T(15-16: 1, 2), B, L*, *Rings, Modules and Linear Algebra.* B. Hartley and T.O. Hawkes. B & N, 1970, 210 pp, $8 (P). A very stiff text with emphasis as in title. The book assumes the student has an excellent background and sophistication. Groups are treated as a special case of modules. Binding impermanent. W.C.R.

ALGEBRA, COMPUTER SCIENCE, L***, T?, *Modern Applied Algebra.* Garrett Birkhoff and Thomas C. Bartee. McGraw, 1970, 431 pp, $11.50. Survey of applications of modern algebra: Boolean algebras, optimization of computer design; lattices; automata theory; communications codes from groups, from polynomial rings, and from finite fields. Does not include enough group theory, etc. to be usable as a text for students without a prior algebra course; did work well as text for a one semester seminar for juniors and seniors following a one year algebra course (skipping the elementary material and the incorrect discussion of ALGOL). Problems and bibliography are good, not excellent. Publisher has problems with misprints and publishing schedules. J.G.L.

ANALYSIS, T(17-18), P, L, *Differential Forms.* Henri Cartan. Houghton Mifflin, 1970, 166 pp, $10.50. Assumes a familiarity with graduate level functional analysis and topology. The book gives a fairly complete treatment of the topic of differential forms with an interesting section on the applications of the moving frame method, and an introduction to the calculus of variations. The effectiveness of the work is greatly weakened by the need to have in hand, for definitions and basic results, the author's previous book, *Differential Calculus.* Although repeatedly referred to by title, no complete bibliographic reference to this needed text is found anywhere in the book. T.A.V.

APPLIED MATHEMATICS, STOCHASTIC CONTROL THEORY, T(18), P, L, *Introduction to Stochastic Control Theory.* Karl J. Åström. Acad Pr, 1970, 299 pp, $14.50. "The purpose of the book is to present theory for analysis, parametric optimization, and optimal control of stochastic control processes." Prior to this a concise introduction to stochastic control, stochastic processes and stochastic models is presented. Annotated bibliography after each chapter. Exercises. R.S.K.

APPROXIMATE QUADRATURE, L. *Quadrature Formulae*. A. Ghizzetti and A. Ossicini. Acad Pr, 1970, 192 pp, $10. Systematic development of quadrature formulas using the tool of integration by parts (i.e. the Green-Lagrange identity relative to a linear differential operator and its adjoint). The method gives an expression for both the integral approximation and for the remainder. It is said to yield all of the classical formulas, many of which are developed. Probably not suitable as a text. R.B.K.

APPROXIMATION THEORY, T(18: 1), S, P, L. *The Functional Method and its Applications: Translations of Mathematical Monographs, Volume 28*. E.V. Voronovskaja. Transl: R.P. Boas. AMS, 1970, 203 pp, $17.80. This book is an English translation of a Russian work. The content lies in the field of approximation theory, the author being somewhat attentive to the needs of the engineer and scientist in his exposition. The topics treated vary from what one might call the usual topics in a work on introductory approximation theory, giving the book a distinctive character. R.J.

APPROXIMATIONS AND EXPANSIONS, *Integrals and Sums: Some New Formulae for their Numerical Evaluation*. P.C. Chakravarti. Oxford U Pr for Athlone Pr, 1970, 88 pp, $7.25. Derivation and generalizations of the Euler-MacLaurin summation formulae using integration by parts, the theory of residues, and operational calculus. R.B.K.

AUTOMATA THEORY, COMPUTERS, LOGIC, L***, *Essays on Cellular Automata*. Ed: Arthur W. Burks. U of Ill Pr, 1970, 375 pp, $12.50. Extensions of Von Neumann's work on self-describing and self-reproducing machines, especially cellular automata; also heuristic uses of computers. By S. Ulam, A.W. Burks, E.F. Moore, R.G. Schrandt, P.R. Stein, J.H. Holland, J.W. Thatcher and J. Myhill. Of interest both to professionals in automata theory and laymen who have been intrigued by the cellular automation "Life" (*Sci. Am.*, Oct., Nov. '70, Feb. '71). A companion volume to Von Neumann's *Theory of Self-Reproducing Automata*. J.G.L.

CALCULUS, T(13: 2), *Calculus with Analytic Geometry*. Nathan O. Niles and George E. Haborak. P-H, 1971, 589 pp, $11.95. Presents calculus as it was done 50 years ago. A working knowledge is provided as claimed, but the further claim--that this is done without sacrificing mathematical validity--is doubtful, especially in the treatment of limits and of the definite integral. K.W.

CALCULUS, T(14-15: 2), *Advanced Calculus*. Hugo Rossi. Benjamin, 1970, 732 pp, $17.50. An advanced calculus approach to "bridge the nebulous gap" between approximately the first three semesters of calculus and modern analysis with applications in science and engineering. Contains a thorough review of linear algebra, a well-done summary of a first course in calculus, then proceeds with differential equations, Fourier Series, and finally the various versions of Stokes' theorem and ends with Dirichlet's principle. A "well done mixture of mathematical rigor and physical intuition." L.L.K.

COMBINATORICS, *Combinatorial Theory and its Applications: Volume I, II and III. Colloquia Mathematica Societatis János Bolyai, 4*. Ed: P. Erdös, A. Rényi and Vera T. Sós. North-Holland, 1970, 1201 pp. Three volumes of papers presented at the Colloquium on Combinatorial

Theory and Its Applications, organized by the Bolyai Mathematical Society. The papers were presented in Balatonfüred, Hungary, in the fall of 1969. R.J.

COMMUTATIVE ALGEBRA, T(18: 2), P, L. *Commutative Algebra.* Hideyuki Matsumura. Benjamin, 1970, 262 pp, $7.95 (P). The basic concepts of flatness, dimension, depth, completion, normal and regular rings are developed with virtually no motivation. The final chapters are on derivations, formal smoothness, Nagata rings and excellent rings. The style is cold and sterile. Assumes a knowledge of homological algebra (Ext and Tor). A scattering of examples and exercises. L.C.L.

COMPLEX ANALYSIS, T?, P, S, L. *Complex Variables.* Robert Ash. Acad Pr, 1971, 255 pp, $9.50. Purporting to be a textbook for the graduate level complex variables course, assuming only an undergraduate course in the field and elementary point set topology, this book gives a list of definitions, lemmas, theorems, proofs and problems common to such a course. Unfortunately, there is virtually no descriptive material included. Although possibly very useful for the instructor of such a course, as a text it fails to give any motivation to the student and completely ignores the beauty and import of complex function theory. T.A.V.

CONFORMAL MAPPING, T(18), P, L. *A Conformal Mapping Technique for Infinitely Connected Regions. Memoirs of the American Mathematical Society, Number 91.* Maynard G. Arsove and Guy Johnson, Jr.. AMS, 1970, 56 pp, $1.80 (P). "Methods of classical analysis devised originally for the disc are here extended to more general plane regions by the use of Green's lines, the Green's mappings, and an ideal boundary structure generalizing the prime-end structure of Carathéodory." R.B.K.

DIFFERENTIAL AND INTEGRAL EQUATIONS, T*(16-17: 1), L*. *Principles of Differential and Integral Equations.* C. Corduneanu. Allyn and Bacon, 1971, 201 pp, $12.25. Considers the theory of differential and integral equations, with the goal of preparing the reader for graduate-level courses on advanced topics in these subjects. Discussion of ordinary differential equations covers existence theory, global problems, linear systems, and stability theory; the book then considers Volterra, Fredholm, and the theory of self-adjoint integral equations. Well written, with challenging problems and many references. Prerequisites: advanced calculus, basic linear algebra, acquaintance with entire and meromorphic functions. D.F.A.

DIFFERENTIAL EQUATIONS, T***(15: 1, 2), B, L. *Ordinary Differential Equations.* William T. Reid. Wiley, 1971, 553 pp, $17.50. Emphasis is on basic existence theorems, two-point boundary problems, and an introduction to the stability and asymptotic behavior of solutions. The mathematical prerequisites are minimal and appendices contain all the necessary matrix theory. The level of rigor is quite high and would enable students to go on to fields of extremals directly. There are many references and few applications. W.C.R.

DIFFERENTIAL EQUATIONS-LINEAR ALGEBRA, T*(14: 1). *Linear Mathematics: An Introduction to Linear Algebra and Linear Differential Equations.* Fred Brauer, John A. Nohel, and Hans Schneider. Benjamin, 1970, 347 pp, $13.95. An introductory chapter gives examples leading to linear systems of equations both algebraic and differential. Chapters on matrices, solution of equations, and determinants lead into the theorems on basis and dimension. The theory is exploited to show similarities between algebraic and differential systems. Chapters on eigenvalues and the Laplace transform finish the book. Text intended for fourth-semester calculus. Exercises (some with solutions) look good. L.A.S.

EDUCATION, E, S, L. *The Effects of Structural Relations on Transfer: Psychological Monographs on Cognitive Processes, Volume II.* Z.P. Dienes and M.A. Jeeves. Humanities Pr, 1970, 148 pp, $6. Raises some interesting questions for those teaching mathematics (or mathematics education) about the extent to which "motivating" concepts by giving simpler versions first interferes with efficient learning. However, that the structures studied are groups appears to be incidental. L.A.S.

EDUCATION, PROBABILITY AND STATISTICS, E, P*, L*. *The Teaching of Probability and Statistics. Proceedings of the First CSMP International Conference Co-sponsored by Southern Illinois University and Central Midwestern Regional Educational Laboratory.* Ed: Lennert Råde. Wiley, 1970, 373 pp, $16. Contains all papers presented at the March, 1969, CMSP (Comprehensive School Mathematics Program) International Conference. Subject of the conference was the teaching of probability and statistics at the pre-college level, but the book will be of interest to college teachers as well. Contains many fascinating articles by acknowledged experts on what is and what can be done at this level (primarily secondary school), and abounds with illustrative problems and examples. Good international bibliography. R.S.K.

FOURIER ANALYSIS, T(18: 1), P, L. *Almost-Periodic Functions and Functional Equations.* Luigi Amerio and Giovanni Prouse. Van Nostrand, 1971, 184 pp, $13.95. In Part I of this book, the authors treat the theory of almost-periodic functions with values in a Banach space; topics presented include harmonic analysis of almost-periodic and weakly almost-periodic functions, and the integration of almost-periodic functions. In Part II, applications to the wave equation, Schrödinger-type equations, and certain nonlinear equations are considered. The book is in the University Series in Higher Mathematics and assumes background in functional analysis. D.F.A.

GENERAL, S, L. *Bourbaki: Towards a Philosophy of Modern Mathematics I.* J. Fang. Paideia Pr, 1970, 144 pp, $5.90. *Hilbert: Towards a Philosophy of Modern Mathematics II.* J. Fang. Paideia Pr, 1970, 205 pp, $6.90. The first and second of six projected volumes whose aim is to present a philosophy of mathematics based on the contents of modern mathematics. Statements and opinions are supported by a liberal dosage of quotations from "working mathematicians" in much the same way as theorems are supported by proofs. The result is a rigid style and a lack of focus (not all of these quotations were directed to the same audience). The reader comes away feeling coerced rather than convinced. Naturally

Bourbaki includes the standard anecdotes and *Hilbert* includes a description and progress report (up to 1969) of the twenty-three problems. L.C.L.

GENERAL, S, P, L*, *The Nature and Growth of Modern Mathematics.* Edna E. Kramer. Hawthorn Books, 1970, 758 pp, $24.95. The mathematical content consists roughly of topics found in the union of other popular mathematical writings. Mathematics since 1950 is not covered as completely as in the COSRIMS report. However, the unique value of the book is its richness. Historical perspectives, anecdotes of human interest, biographical sketches are woven into each concept and provide identity, continuity and inspiration of special value to students receiving a rapid, concentrated exposure to advanced topics. Includes a large index; no exercises. L.C.L.

GENERAL, S, B, L, *Adventures Among the Toroids.* B.M. Stewart. Published by the author, 4494 Wausau Road, Okemos, Michigan 48864. 1970, 206 pp, $6 (P). Investigations of the 2-dimensional polyhedra (oriented, unbounded, and with regular faces) which are spheres with handles. Many drawings, and instructions for models. Lots of projects, elementary to quite deep, outlined for self-study. Size (5" x 13") is not successful; printing (hand-lettered) is. Unique, and even with the limitations confessed by the author, fascinating. L.A.S.

GENERAL TOPOLOGY, T***(16-17: 2), L*, *General Topology.* Stephen Willard. A-W, 1968, 369 pp, $13.50. A strong and careful exposition of general topology which serves as a text and reference. The book is written as a text for good undergraduate or first-year graduate students. It has a substantial section on historical notes, and an extensive bibliography and index. In a table of subject dependence, the author suggests several orders in which the text topics can be studied. R.J.

*HISTORY, FOUNDATIONS OF ANALYSIS, T(15-16: 1), S, B, L, *The Development of the Foundations of Mathematical Analysis from Euler to Riemann.* I. Grattan-Guinness. MIT Pr, 1970, 186 pp, $10. Exactly what the title indicates. An historical development of the foundations of the basic concepts in mathematical analysis: theories of functions and continuity, limits, convergence of series (including Fourier series), as well as the formation of the derivative and integral. The theme is developed around the solution of the vibrating string problem. The aim is "not to give a detailed account of the discussion itself, but to show how it was influenced by, and then influenced, the development of the foundations of analysis of its time." Recommended as resource material for a reading course or seminar for post-advanced calculus students. Also, teachers of analysis should take note. R.B.K.

HYPERBOLIC MANIFOLDS, HOLOMORPHIC MAPPINGS, T(18: 1, 2), P, L, *Hyperbolic Manifolds and Holomorphic Mappings.* Shoshichi Kobayashi. Marcel Dekker, 1970, 148 pp, $11.75. Based on lectures at Berkeley 1969-70. An extremal pseudodistance complementary to Carathéodory's is introduced on every complex manifold. The big Picard Theorem, viewed as an extension theorem for holomorphic mappings, is developed and applied. Author and subject index included. R.B.K.

HYPERBOLIC MAPPINGS, T(18), P, L. *Intrinsic Measures on Complex Manifolds and Holomorphic Mappings. Memoirs of the American Mathematical Society, Number 96.* Donald A. Eisenman. AMS, 1970, 80 pp, \$1.90 (P). From the Poincaré-Bergman metric, an invariant measure is defined on the unit ball B^n in C^n and generalized to general complex manifolds. Applications to holomorphic mappings. R.B.K.

NUMBER THEORY, T*(17: 1), P, B, L*. *Algebraic Theory of Numbers.* Pierre Samuel. Transl: Allan J. Silberger. Houghton Mifflin, 1970, 109 pp, \$7.95. A very deep and very compact treatment of algebraic number theory, this book assumes a solid algebra course. In addition, an elementary number theory course would help. A modest number of problems and few of a routine nature. People considering Pollard should consider this text. W.C.R.

NUMERICAL ANALYSIS, DIFFERENTIAL EQUATIONS, COMPUTERS, PHYSICS, T?, L. *Introduction to Numerical Analysis and Applications.* Donald Greenspan. Markham, 1971, 182 pp, \$9.50. "Numerical Analysis Applied to Physics." One semester undergrad text heavily weighted towards differential equations. Includes chapters on: Exact solutions of ODE's (for the underprivileged) , a discrete model of physics (which belongs in an innovative physics course), and interval arithmetic error analysis. Slights many standard and important topics--not suitable as a text for a general audience. Good bibliography, references. J.G.L.

NUMERICAL METHODS, T(18: 1), P, L. *The Approximate Minimization of Functionals.* James W. Daniel. P-H, 1971, 228 pp, \$9.50. In its opening chapters, this book considers variational problems in an abstract setting, and theory and examples of discretization. Discussion of gradient-type Banach space methods and conjugate-gradient methods in Hilbert space follow, and the book concludes with the development of computational methods of minimization (primarily for unconstrained problems) in real Euclidean spaces. Contains about 100 problems (of a theoretical nature) and 200 references. D.F.A.

OPTIMIZATION, T(14-16: 1), S, L. *Optimization Theory.* David Russell. Benjamin, 1970, 405 pp, \$17.50. Typescript lecture notes. Fundamentals from advanced calculus and linear algebra are contained where needed. Covers block search, least squares, Newton's, gradient and gradient project methods and linear equality and inequality constraints. All confined to R^n. No references. No index. R.W.N.

PHYSICS, P, L. *An Introduction to Current Algebra.* D.H. Lyth. Oxford U Pr, 1970, 67 pp, \$3 (P). An introduction to the concept of the algebra of currents concluding with a chapter on the future of current algebra applications. Possibly a source for a mathematician interested in this new structure in physics. Requires postgraduate knowledge of elementary particle theory. J.A.S.

REAL ANALYSIS, T(17-18: 2). *Analyse: Topologie générale et analyse fonctionnelle.* Laurent Schwartz. Hermann, 1970, 432 pp, 58F. An extended development of *Cours d'analyse*, Hermann, 1967, by the same author. No exercises and no bibliography. This is intended as the second part of a two volume treatise on analysis. Very Bourbaki! J.A.S.

REMEDIAL, T*(13: 1), *Intermediate Algebra.* Ward D. Bouwsma. Macmillan, 1971, 347 pp, $8.50. Covers the field axioms, linear equations and inequalities, quadratic equations, systems of equations, complex numbers, polynomial, exponential and logarithmic functions and finite modular systems. Attractively printed, it includes interesting applications and historical notes. Exercises are incorporated. J.N.C.

REMEDIAL, T(13: 2), *Algebra and Trigonometry.* Marvin Marcus and Henryk Minc. Houghton-Mifflin, 1970, 500 pp, $9.95; and *College Algebra.* Marvin Marcus and Henryk Minc. Houghton-Mifflin, 1970, 374 pp, $8.95. The two texts differ only in the chapter on trigonometry added to the one which is quite adequate. This reviewer did speculate on why the winding function was not used since the necessary background in coordinates was available. These texts may prove a bit sophisticated as to notation for students with but one year of high school algebra. The chapter on matrices is quite readable. However, determinants and vectors are merely mentioned. C.S.C.

REMEDIAL, T(13: 1), *Preparation for Calculus.* William L. Hart. Intext Educational Pub, 1971, 367 pp, $8.95. Another text designed to fit the recommendations of the CUPM pre-calculus course Mathematics O. It is a classical approach and looks like any 12th grade course in pre-calculus. A student could use it as a supplement to a calculus course to fill in gaps in his mathematics background. The text begins with a very elementary review, presents a foundation for plane analytic geometry, standard trigonometric functions, exponential and logarithm functions, inequalities, induction, and polynomial functions. L.L.K.

SOCIAL SCIENCES, S, B, L, *Mathematical Sciences and Social Sciences.* Ed: William Kruskal. P-H, 1970, 83 pp, $4.95. Members of the Mathematical Sciences Panel of the Behavioral and Social Sciences Survey (1967-69) have provided a series of essays illuminating the uses of mathematics, statistics, and computation in the social sciences. Discussed among other topics are cooperation between mathematicians and social scientists both in research and the teaching of mathematics to social-science students, social and technical problems in data collecting, and the possible consequences of these problems in (mis-)guiding public policy. L.A.S.

STATISTICS, T(15-16: 1, 2), *Introductory Mathematical Statistics.* Erwin Kreyszig. Wiley, 1970, 470 pp, $12.50. Presupposes calculus. Includes chapters on quality control, error analysis, and decision functions aside from the standard topics. No discussion of subjective probability. F.L.W.

STATISTICS, S, P, L, *Reliability Handbook.* B.A. Kozlov and I.A. Ushakov. Transl: Lisa Rosenblatt. HR & W, 1970, 391 pp, $9.50. "Designed as a practical manual for everyday work of ensuring reliability of ... various complicated systems." Assumes a "full knowledge of reliability theory." F.L.W.

STATISTICS, T?, S, *Statistical Reasoning.* Lloyd Rosenberg. Merrill, 1971, 94 pp, $3.95, $1.95 (P). Presents basic information on point and interval estimation, hypothesis testing and decision theory in a somewhat limited fashion, assuming a background of

elementary calculus and probability theory. Topics are developed as solutions to stated problems. Would have to be combined with other material or augmented to be used effectively as a text, but can be read independently as an introduction to these concepts. R.S.K.

STATISTICS, TABLES, P, L, *Selected Tables in Mathematical Statistics, Volume I*. H.L. Harter and D.B. Owen. Markham, 1970, 405 pp, $5.80. First of a possible series of statistical tables sponsored by the Institute of Mathematical Statistics. Contains four new sets of tables plus an extension of the tables of critical values and probability levels for the Wilcoxon Rank Sum Test and the Wilcoxon Signed Rank Test. Photo-offset printing. R.S.K.

TOPOLOGY, T*(15-16: 2), S*, *Undergraduate Topology*. Robert H. Kasriel. Saunders, 1971, 285 pp, $10.50. Very much a beginning text relying more on analysis (calculus) than modern algebra. Over half the book is naive set theory (without ordinals) and metric spaces. The general topology part continues the orientation to analysis as typified, for example, by the section on nets. Paracompactness is not even defined. Many nice diagrams and a reasonable number of standard examples. Good clear writing and numerous elementary exercises. Could be used as an excellent supplement to advanced calculus. J.A.S.

TOPOLOGY, S, P, L, *Proceedings of the Washington State University Conference on General Topology*. Pi Mu Epsalon Society. 1970, 136 pp, $3 (P). Sixteen papers presented at Washington State University in March 1970. Thirteen of the papers are brief, some expository. Contains 3 longer papers by Z. Frolík, E.A. Michael and A.H. Stone. T.A.V.

Reviewers Whose Initials Appear Above

David F. Appleyard, Carleton; Clarence S. Carlson, St. Olaf; Judith N. Cederberg, St. Olaf; Richard Jarvinen, Carleton; Lorraine L. Keller, St. Olaf; Roger B. Kirchner, Carleton; Richard S. Kleber, St. Olaf; Loren C. Larson, St. Olaf; John G. Lewis, St. Olaf; R.W. Nau, Carleton; William C. Ramaley, Carleton; J. Arthur Seebach, Jr., St. Olaf; Linda A. Seebach, St. Olaf; T.A. Vessey, St. Olaf; Kenneth Wegner, Carleton; Frank L. Wolf, Carleton.

CORRECTIONS

SEVERAL COMPLEX VARIABLES, *Theorie der Funktionen mehrerer komplexer Veränderlichen*. H. Behnke and P. Thullen. Springer-Verlag, 1970, telegraphic review, March 1971, is *Volume 51* of *Ergebnisse der Mathematik und ihrer Grenzbebiete*.

NEWS AND NOTICES

EDITED BY RAOUL HAILPERN, SUNY at Buffalo

Readers are invited to contribute to the general interest of this Department by sending news items to Mathematical Association of America, 1225 Connecticut Avenue, N.W., Washington, D.C. 20036. Items must be submitted at least two months before publication can take place.

PERSONAL ITEMS

Professor K. G. Clemans, Southern Illinois University, represented the Association at the inauguration of O. R. Herron as president of Greenville College on November 14, 1970.

Professor Raymond Killgrove, California State College, Los Angeles, represented the Association at the inauguration of F. M. Binder as president of Whittier College on November 5, 1970.

Professor Emeritus V. V. Latshaw, Lehigh University, represented the Association at the inauguration of Dr. Michelini as president of Wilkes College on November 21, 1970.

Professor M. F. Smiley, State University of New York at Albany, represented the Association at the inauguration of Reverend M. T. Conlin, O.F.M., as president of Siena College on October 24, 1970.

Brooklyn College, City University of New York: Assistant Professor L. L. Gavurin has been promoted to Associate Professor; Associate Professor Meyer Jordan has been promoted to Professor.

California State College, Fullerton: Drs. J. J. Bucuzzo, University of Notre Dame, and H. S. Shultz, Purdue University, have been appointed Assistant Professors; Associate Professor R. V. Benson has been promoted to Professor.

California State College, Long Beach: Associate Professors W. D. James and N. E. Sexauer have been promoted to Professors; Assistant Professor A. P. Gittleman has been promoted to Associate Professor.

California State Polytechnic College, Kellogg-Voorhis: Dr. Sidney Birnbaum, University of South Carolina, has been appointed Professor and Chairman of the Mathematics Department; Assistant Professor Thomas Flynn retired in June 1970 with the title of Professor Emeritus.

University of California, Los Angeles: Assistant Professors Nathaniel Grossman and David Sanchez have been promoted to Associate Professors.

Cleveland State University: Assistant Professor Nosup Kwak, Brooklyn College of the City University of New York, has been appointed Assistant Professor; Assistant Professors J. N. Hanson and R. L. Pruitt have been promoted to Associate Professors.

Colgate University: Drs. Richard Ringeisen, Michigan State University, and Larry Shatoff, Duke University, have been appointed Assistant Professors.

Concordia College: Mr. James Forde, Vanderbilt University, has been appointed Assistant Professor; Associate Professor C. V. Heuer has been appointed Chairman of the Mathematics Department.

Fort Lewis College: Assistant Professor L. S. Johnson, Western Illinois University, has been appointed Assistant Professor; Mr. H. E. Stocker has been promoted to Assistant Professor.

Illinois State University: Drs. S. H. Friedberg, CUPM, J. T. Parr, University of Illinois, and Lawrence Spence, Michigan State University, have been appointed Assistant Professors.

Iowa State University: Dr. B. E. Cain, MIT, has been appointed Visiting Assistant Professor; Associate Professors J. L. Cornette, J. A. Dyer and J. C. Mathews have been promoted to Professors.

Kansas State University: Assistant Professor Gregory Bachelis, SUNY at Stony Brook, has been appointed Associate Professor; Dr. L. M. Chawla, Sargodha, West Pakistan, has been appointed Professor; Drs. W. D. Curtis, University of Massachusetts, Robert Dressler, Southern Illinois University, L. M. Herman, Plymouth State College, and Janina L. Spears, Kansas State University, have been appointed Assistant Professors; Assistant Professor R. J. Greechie has been promoted to Associate Professor.

Kent State University: Associate Professor Heinz Renggli, University of Wisconsin, has been appointed Professor; Assistant Professor F. L. Sandomierski, University of Wisconsin, has been appointed Associate Professor; Mr. Richard Little has been promoted to Assistant Professor; Professor L. E. Bush retired on September 15, 1970 with the title of Professor Emeritus.

Kutztown State College: Mr. W. E. Bateman has been promoted to Assistant Professor; Assistant Professor J. C. Gerhard has been promoted to Associate Professor; Professor J. D. Daugherty retired in June 1970 with the title of Emeritus Professor.

Loyola University, Chicago: Dr. M. G. Buntinas, Illinois Institute of Technology, has been appointed Assistant Professor; J. R. VandeVelde, S. J., has been appointed Chairman of the Mathematics Department.

Mankato State College: Assistant Professors James Andersen and James Pelzl have been promoted to Associate Professors.

Marietta College: Associate Professor Neil Bernstein, SUNY at Fredonia, has been appointed Associate Professor; Dr. R. H. Pitasky, Rutgers University, has been appointed Assistant Professor; Professor Theodore Bennett retired in May 1970.

Mohawk Valley Community College: Dr. James Burns, Indiana University, has been appointed Assistant Professor; Messrs. Harold Fitzpatrick and Norbert Oldani have been promoted to Assistant Professors; Associate Professor L. A. Trivieri has been promoted to Professor.

Nicholls State University: Dr. L. S. Haw, Clemson University, has been appointed Associate Professor and Head of the Mathematics Department; Dr. M. M. Ohmer is now Professor of Mathematics and Dean of the College of Sciences.

North Carolina State University: Assistant Professor W. G. Dotson, Jr., has been promoted to Associate Professor; Mr. C. N. Anderson has been promoted to Assistant Professor.

Old Dominion University: Dr. G. W. Pfeiffer, University of Georgia, has been appointed Assistant Professor; Dr. J. J. Swetits, Lafayette College, has been appointed Associate Professor; Professor H. S. Grant retired on June 1, 1970.

Pacific Lutheran University: Drs. J. E. Brink, Iowa State University, and N. C. Meyer, Jr., University of Oregon, have been appointed Assistant Professors; Mr. R. S. Fisk has been promoted to Assistant Professor.

University of Puerto Rico: Mr. J. P. Morales has been appointed Instructor; Associate Professor Dorothy Bollman has been promoted to Professor.

Rutgers University: Dr. David Dobbs, UCLA, has been appointed Assistant Professor; Associate Professor Tilla K. Milnor, Boston College, has been appointed Professor; Professor Francois Treves, Purdue University, has been appointed Professor; Associate Professor Benjamin Muckenhoupt has been promoted to Professor.

St. John's University: Dr. Donald McCarthy, Fordham University, has been appointed Assistant Professor; Associate Professor Michael Capobianco, Chairman of the Mathematics Department, has been promoted to Professor.

University of Southwestern Louisiana: Dr. Duane Blumberg, University of Wisconsin, has been appointed Assistant Professor; Assistant Professor Steve Ligh, University of Florida, has been appointed Associate Professor; Assistant Professor Henry Heatherly has been promoted to Associate Professor.

State University College at Brockport: Assistant Professor Norman Bloch, Wayne State University, has been appointed Assistant Professor; Assistant Professor Theron

Rockhill has been promoted to Associate Professor.

Stephen F. Austin State University: Assistant Professor C. W. Proctor, Memphis State University, has been appointed Assistant Professor; Mr. J. T. Robbins has been promoted to Assistant Professor.

University of Utah: Dr. L. G. Lewis, CUNY, has been appointed Assistant Professor; Assistant Professor Klaus Schmitt has been promoted to Associate Professor.

Virginia Commonwealth University: Dr. E. A. Newburg, University of Illinois, has been appointed Associate Professor; Drs. R. J. Schwabauer, Loyola University, Louisiana, W. A. Thedford, New Mexico State University, and Mr. D. G. Wilson, University of Maryland, have been appointed Assistant Professors.

Virginia Polytechnic Institute: Drs. L. C. Baird, Indiana University, M. B. Boisen, University of Nebraska, P. F. Duvall, U. S. Army, and W. R. Woodward, University of Nebraska, have been appointed Assistant Professors; Dr. C. W. Patty has been appointed Head of the Mathematics Department; Assistant Professor L. S. Husch has been promoted to Associate Professor.

Associate Professor C. R. Deeter, Texas Christian University, has been promoted to Professor.

Associate Professor J. A. Ernest, University of California, Santa Barbara, has been promoted to Professor.

Mr. Gordon Gregersen, Eastern Oregon College, has been appointed Assistant Professor at Carroll College.

Associate Professor L. A. Hart, Chairman of the Department of Mathematics at Loras College, has been promoted to Professor.

Brother Thomas Kiesler, CFX, Bellarmine College, has been appointed Associate Professor at Spalding College.

Dr. C. S. Kim, University of Oklahoma, has been appointed Assistant Professor at Indiana University, Southeast.

Assistant Professor R. E. Mullins, Marquette University, has been promoted to Associate Professor.

Dr. F. R. Norris, Vanderbilt University, has been appointed Lecturer at the University of North Carolina, Wilmington.

Miss Frances C. Pascale, Albertus Magnus College, has been promoted to Assistant Professor.

Assistant Professor Robert Pendleton, Louisiana State University, has been appointed Associate Professor and Chairman of the Mathematics Department at Whittier College.

Dr. E. M. Scheuer, RAND Corporation and C-E-I-R Professional Services Division of Control Data Corporation, has been appointed Associate Professor of Management Science at San Fernando Valley State College.

Dr. W. E. Shreve, University of Connecticut, has been appointed Assistant Professor at North Dakota State University.

Associate Professor J. L. Smith, Middle Tennessee State University, has been promoted to Professor.

Dr. C. W. Snyder, University of Denver, has been appointed Assistant Professor at St. Lawrence University.

Associate Professor D. H. Staley, Ohio Wesleyan University, has been promoted to Professor.

Associate Professor J. H. Stoddard, Kenyon College, has been appointed Professor and Chairman of the Department of Mathematics at Upsala College.

Mrs. Ida Arbeit Sussman, Adelphi University, has been promoted to Assistant Professor.

Assistant Professor J. T. Teska, Ripon College, has been promoted to Associate Professor.

Dr. Mary E. Von Wolff, St. Norbert's College, has been appointed Associate Professor at the University of South Alabama.

Assistant Professor R. J. Weaver, University of Massachusetts, has been appointed Assistant Professor at Mount Holyoke College.

Assistant Professor Clifton Whyburn, Louisiana State University, has been appointed Associate Professor at the University of Houston after spending the 1969–70 year in Germany on an Alexander von Humboldt Fellowship.

THE MATHEMATICAL ASSOCIATION OF THE UNITED KINGDOM

The Mathematical Association of the U. K. is celebrating its centenary this year and to mark the occasion is publishing a special "Centenary Issue" of its journal, THE MATHEMATICAL GAZETTE. The articles in this special issue are, in the main, reprints of material published in earlier numbers, which have been selected to show the extent and variety of the work of the Association. The origins of the Association and the development of its policies are described, and a section of the issue is devoted to articles and addresses by some of the Association's more noted past members. These articles include *The teaching of Euclid* by Bertrand Russell, *The integral* $\int_0^\infty (\sin x\, dx)/x$ by G. H. Hardy, and *Address on relativity* by Sir Arthur Eddington. A collection of smaller papers includes *Some arithmetical puzzles* by W. W. Rouse Ball and some early reviews, including E. T. Whittaker on *Lamb's Calculus* and one on *Bulletin No. 1* of the Association of Teachers of Mathematics in the Middle States and Maryland.

In view of the interest that this issue is likely to arouse, the MAA has arranged for its members to purchase copies directly from the Washington office (1225 Connecticut Avenue, N. W., Washington, D. C. 20036) at a price of $2.50 postpaid.

MATHEMATICAL ASSOCIATION OF AMERICA

Official Reports and Communications

THE FIFTY-FOURTH ANNUAL MEETING OF THE ASSOCIATION

The Fifty-fourth Annual Meeting of the Mathematical Association of America was held at the Chalfonte-Haddon Hall, Atlantic City, New Jersey, from Saturday to Monday, January 23 to 25, 1971, in conjunction with meetings of the American Mathematical Society, the National Council of Teachers of Mathematics, and the Association for Symbolic Logic. The sessions of the Association on Saturday and Sunday were joint meetings with the National Council of Teachers of Mathematics. There were registered 3513 persons, including 2003 members of the Association.

Sessions of the Association were held on Saturday morning, Sunday morning, and Monday morning and afternoon in the Pennsylvania Room of Haddon Hall. Presiding officers were Professor F. E. Clark at the first two lectures on Saturday morning, Professor L. G. Woodby at the panel discussion on Saturday morning, Dean J. N. Eastham at the first lecture on Sunday morning, and Dr. J. B. Kruskal at the presentations on "How to Write Mathematics," Professor Emil Grosswald on Monday morning, Professor Albert Wilansky at the presentation of the Report on the 1970 International Congress of Mathematicians in Nice on Monday afternoon, and Professor H. F. Trotter at the last lecture on Monday afternoon.

The Program Committee consisted of Joshua Barlaz, Chairman; J. N. Eastham, Emil Grosswald, J. B. Kruskal, H. F. Trotter, Albert Wilansky, and L. G. Woodby.

FIRST SESSION OF THE ASSOCIATION

Joint Session with the National Council of Teachers of Mathematics

The Preparation of Mathematics Teachers

The Preparation of Mathematics Teachers in the United States During the Past Twenty Years, by Dr. C. R. Phelps, Conference Board of the Mathematical Sciences

The restructuring of the training of mathematics teachers is closely linked with two major movements—institutes for teachers and curricular reforms. Institutes have provided 175,000 training opportunities, involving over half of the secondary school teachers of mathematics, many leaders in elementary school mathematics, and several thousand senior mathematics faculty. Substantial efforts in alleviating training gaps of many teachers originally trained in other fields should continue since the supply of certificated teachers is still below the demand. Academic year institutes and sequential summer institutes have devised comprehensive programs emphasizing subject matter oriented toward teaching needs. These have graduated about ten thousand teachers with master's degrees in the teaching of mathematics; of these about a quarter have transferred to junior colleges where they form half the mathematics faculty. Concurrently, large curricular reform programs have created strong incentives for further study, which must be oriented toward broad preparation for continuing potential reforms. Institute experiences and curricular reforms are both reflected in the influential CUPM recommendations for teacher preparation and certification.

The Program for the Preparation of Mathematics Teachers in Denmark, by Professor Bent Christiansen, The Royal Danish School of Educational Studies

The Danish school comprises *Folkeskolen* (grade 1–9 (10)) and *Gymnasiet* (grade 10–12). The teaching in G presupposes a masters degree, while the teachers of F receive a $3\frac{1}{2}$-year education at colleges building mainly on the twelve year school. Thus the background of a mathematics teacher in grade 8–10 is the curriculum of *Gymnasiet* and 26 semester hours of mathematics chosen as a major topic at the college. He may, however, receive further mathematics knowledge by participation in the in-service program under The Royal School. The speaker gave an outline of this system emphasizing the new demands on teachers training caused by the changing goals for school-mathematics. Especially he pointed to the increasing role of the didactics of mathematics—taught by *mathematicians*—as a part of the education of teachers.

Panel Discussion: Trends for the Future in the Preparation of Mathematics Teachers

A panel discussion with Professor E. G. Begle, Stanford University, Dr. J. H. Hlavaty, Past President, National Council of Teachers of Mathematics, and Professor Izaak Wirszup, University of Chicago, moderated by Professor L. G. Woodby, Michigan State University.

Professor Begle observed that, in the past half-century, a vast number of studies on teacher characteristics and teacher behaviors have been carried out. This literature should be examined before programs for the training of teachers are developed. Empirical evidence shows that many of our popular beliefs about teaching are myths and hence need not be built into teacher training programs.

Dr. Hlavaty suggested that, in view of changing relations among schools, school boards and communities, and teachers and teacher organizations, the preparation, certification, and continued in-service education of teachers can and should become a responsibility of teacher organizations. Both the professional organizations and the bargaining groups have a vital role. Colleges, universities, and other teacher-training institutions should review and re-design their programs in cooperation with teacher organizations. Prescriptions for pre-service education, realistic apprentice-type experiences for future teachers with experienced in-service teachers, statements of requirements for teacher-certification, and organized plans for continued in-service education should be reached by cooperative consultation between the teacher-training institutions and the teacher-organizations.

Professor Wirszup first discussed some recent reforms in mathematics education in the Soviet Union: 1) The new school mathematics curriculum (1968); 2) The new program for training mathe-

matics teachers (1970); 3) Secondary schools for the mathematically talented; 4) Research in the psychology of learning and teaching mathematics; 5) The teaching of mathematics to handicapped children.

Suggestions for the United States were then made concerning the following: 1) The elementary-school mathematics curriculum; 2) The mathematical training of elementary-school teachers and the need for specialized mathematics teachers from the fourth grade on; 3) Mathematics and education courses in teacher-training programs; 4) Mathematical instructions for (a) disadvantaged children, (b) slow learners, (c) the handicapped, and (d) the mathematically gifted.

SECOND SESSION OF THE ASSOCIATION

Joint Session with the National Council of Teachers of Mathematics

The Computer and the Calculus, by Professor W. B. Stenberg, University of Minnesota, Minneapolis

Computing (or more properly—algorithmic) concepts hold promise of playing an important role in the presentation not only of calculus, but also in linear algebra and ordinary differential equations and certain topics in high school and junior high school mathematics. Among the reasons for this are: the dynamic effect of the algorithmic method; the computing attitude toward variables; the fostering of problem solving skills; the relation of algorithm construction to theorem proving; the beneficial experience of "teaching" a machine to solve a problem.

The introduction of computing concepts in the CRICISAM calculus text profoundly affected the manner in which the subject was developed both in the order of topics and the methods of proof. For example: integration introduced before differentiation; uniform continuity before pointwise continuity; algorithmic (constructive) proofs of existence theorems.

A very short course (10 minutes) in computing was presented, followed by samples from the CRICISAM text showing how calculus develops from the algorithmic approach. These were: the Newton square root algorithm motivating the definition of convergence of sequences; the arithmetic-geometric mean algorithm motivating a completeness axiom; algorithms for area leading to the definition of the integral. Finally an algorithm was presented for generating tables of sines and cosines purely from integration theory.

Annual Business Meeting of the Association; the Association's Tenth Award for Distinguished Service to Mathematics; Award of the 1971 Chauvenet Prize.

How to Write Mathematics, presentations by Professor Harley Flanders, Tel Aviv University, Mr. George Fleming, W. B. Saunders Company, Professor P. R. Halmos, Indiana University, and Professor George Piranian, University of Michigan.

Professor Flanders gave an introductory statement, referred to his article *Manual for Monthly Authors* (this MONTHLY, January 1971), and introduced the other speakers.

Professor Halmos discussed the strategy of writing mathematics. He described his own method of writing, and referred to his forthcoming article *How to Write Mathematics* (L'Enseignement Mathématique, 1971).

Professor Piranian spoke about the tactics and nitty gritty of mathematical writing, and he presented the research journal editor's viewpoint.

Mr. Fleming described mathematics writing from the publisher's viewpoint. He presented a picture of cooperation that should exist between the editor and author, and how the publisher's editors can best serve their authors.

THIRD SESSION OF THE ASSOCIATION

Topics in Mathematical Logic—An Introduction

Model Theory, by Professor Abraham Robinson, Yale University

Following an introduction which included some simple direct applications of Model Theory to Algebra, the speaker concentrated on the metamathematical analysis of the concept of an alge-

braically closed field. This leads in a natural way to the notions of model completeness and model completion and their generalizations. Hilbert's seventeenth problem on sums of squares was discussed in this context. Some recent results that were first obtained by the use of forcing were also included.

Set Theory, by Professor J. R. Shoenfield, Duke University

The speaker gave an expository survey of work in axiomatic set theory, especially on the work of the last decade. The chief topics were independence proofs and new axioms for set theory.

Recursion Theory, by Professor G. E. Sacks, Massachusetts Institute of Technology and Yale University

The speaker reviewed forty years of recursion theory: a survey of the key ideas of recursion theory from Gödel (1931) to the present day (1971) with a guess at what the future holds.

FOURTH SESSION OF THE ASSOCIATION

Report on the 1970 International Congress of Mathematicians in Nice, by Professor Meyer Jerison, Purdue University, Professor Daniel Gorenstein, Rutgers, The State University, and Professor C. C. Hsiung, Lehigh University

Professor Jerison described the general program and arrangements for the meeting and reviewed briefly the work of the four Field Medal winners, namely Alan Baker of England, Heusuke Hironaka of Japan, Sergei Novikov of the U.S.S.R., and John Thompson of the USA.

Professor Gorenstein presented a summary of those events related to finite group theory.

Professor Hsiung reported that the differential geometry section (C3) met in two afternoons, during which five 50-minute lectures were given by D. Gromoll, S. Kobayashi, G. Mostow, M. Narasimhan, and A. Pogorelov; one lecture was cancelled due to the absence of the speaker K. Nagami. There were some overlappings among this section C3, the section (C2) "Topology of Varieties" and the section (C4) "Analysis on Varieties"; some lectures in C2 and C4 could also be included in C3. One of the most attractive and important lectures among all 1-hour and 50-minute lectures was the 1-hour lecture given by S. S. Chern entitled "Differential Geometry: its Past and Future." He first commented briefly on the following fundamental developments in differential geometry and its related subjects in the last three decades or so: Lie groups, fiber spaces, variational methods and elliptic differential systems, and then formulated ten interesting problems based on the Riemannian structures of differentiable manifolds. There is no doubt that these problems will play an important role in the future of differential geometry.

The Differentiation of Integrals, by Professor A. M. Bruckner, University of California, Santa Barbara

The notion of the derivative of a function generalizes in a natural way to the notion of a pointwise derivative of an integral (or measure) σ with respect to a measure μ. The speaker motivated this generalization, discussed the extent to which the classical differentiation theory generalizes, and gave examples of a number of areas of mathematics in which such derivatives of measures arise in natural ways. These areas include multiple Fourier series, vector analysis, surface area, boundary behavior of harmonic and analytic functions and topological measure spaces. The theory is in turn generalized to a theory of Functional Differentiation Systems which gives some insights into why certain differentiation theorems have close analogues in other parts of mathematics.

SPECIAL SESSIONS OF THE ASSOCIATION

Film showings were held in the Viking Theater in Haddon Hall on Friday evening and, because of an overflow audience that evening, rescheduled for the Pennsylvania Room on Saturday and Sunday evenings. The following films were shown:

Friday,	7:30– 7:40 P.M.	*Sampler from the Topology Films Project*
	7:40 P.M.	*Films of the MAA Individual Lectures Film Project* (ILFP)
	7:40– 8:05 P.M.	SHAPES OF THE FUTURE I—SOME UNSOLVED PROBLEMS IN GEOMETRY—TWO DIMENSIONS with Victor Klee (in color)
	8:10– 8:50 P.M.	SHAPES OF THE FUTURE II—SOME UNSOLVED PROBLEMS IN GEOMETRY—THREE DIMENSIONS with Victor Klee (in color)
	9:00– 9:13 P.M.	*Another Sampler from the Topology Films Project*
	9:15– 9:40 P.M.	A second showing of the film previously shown at 7:40 P.M.
Saturday,	7:30 P.M.	*Encyclopaedia Britannica Sound Filmstrips*
	7:30 P.M.	INTRODUCTION TO CALCULUS: LIMITS AND FUNCTIONS
	7:30– 7:42 P.M.	IDEAS, NUMBERS, AND LIMITS
	7:44– 7:56 P.M.	MORE LIMITS
	8:05 P.M.	INTRODUCTION TO CALCULUS: SEQUENCES AND CONVERGENCE
	8:05– 8:24 P.M.	IMAGINING SEQUENCES
	8:26– 8:46 P.M.	GETTING DOWN TO TERMS
	9:00 P.M.	*Films of the NCTM Series: Elementary Mathematics for Teachers and Students* (in color)
	9:00– 9:10 P.M.	EQUIVALENT FRACTIONS (WATERMELONS)
	9:12– 9:20 P.M.	EXTENDING MULTIPLICATION TO RATIONAL NUMBERS (CLOUDS)
	9:21– 9:31 P.M.	THE BIGGEST RECTANGLE
	9:32– 9:40 P.M.	EXPLOITATION OF ERRORS (EDGAR'S GUESS)
	9:41– 9:48 P.M.	GAMES
	9:51– 9:58 P.M.	PRODUCT OF TWO NEGATIVE NUMBERS (FISH)
	10:00–10:11 P.M.	PROBABILITY (RAJAH)
Sunday,	7:30– 8:37 P.M.	FIXED POINTS, A Lecture by Solomon Lefschetz (A film from the MAA MATHEMATICS TODAY Series, in color)
	8:50– 9:50 P.M.	NONSTANDARD ANALYSIS with Abraham Robinson (A film of the MAA ILFP, b & w).

MEETING OF THE BOARD OF GOVERNORS

The Board of Governors of the Association met on Friday morning in the West Room of Haddon Hall with 36 members present.

The Board approved the appointment by President Young of the following Nominating Committee for 1971: C. B. Allendoerfer, Chairman; R. C. Buck, and D. B. Goodner.

The Board elected Professor J. W. Jewett Second Vice-President of the Association for the period 1971–72.

Professor H. M. Gehman submitted his resignation as a member of the Finance Committee, effective at the end of the Atlantic City meeting. The Board elected Professor G. S. Young to fill the unexpired part of the term of Professor Gehman, extending through 1973.

Professor David Drasin has requested to be relieved of his position as Associate Editor of the MONTHLY in charge of MATHEMATICAL NOTES and CLASSROOM NOTES, effective May 1, 1971. In accordance with a nomination made by Professor Harley Flanders, Editor, the Board elected Professor Robert Gilmer of Florida State University to fill the unexpired part of Professor Drasin's term, extending through 1973.

Professor Victor Klee has requested to be relieved of his position as Associate Editor in charge of the Research Problem Section of the MONTHLY, effective January 1, 1971. In

accordance with a nomination made by Professor Flanders, the Board elected Professor R. K. Guy of the University of Calgary to fill the unexpired part of Professor Klee's term, extending through 1973.

Professor H. A. Thurston has requested to be relieved of his position as Associate Editor of the MATHEMATICS MAGAZINE, effective January 1, 1971. In accordance with a nomination made by Professor G. N. Wollan, Editor, the Board elected Professor P. J. Zwier of Calvin College to fill the unexpired part of Professor Thurston's term, extending through 1973.

The Board approved the recommendation of the Joint Administrative Committee (of MAA, AMS, and SIAM) that, for the Employment Register, the $5 late registration fee for employers be eliminated and that henceforth a $10 registration fee be charged all employers.

The Board approved the following schedule of future meetings of the Association: Pennsylvania State University, August 30–September 1, 1971; Las Vegas, Nevada, January 19–21, 1972; Dartmouth College, August 28–30, 1972; Dallas, Texas, January 27–29, 1973; University of Montana, August 20–22, 1973; San Francisco, California, January 17–19, 1974; Shoreham Hotel, Washington, D. C., January 25–27, 1975; San Antonio, Texas, January 24–26, 1976.

The Executive Director reported the membership of the Association as 18,713 individual members, an increase of 361 over the corresponding date last year, 3 corporate members, and 324 academic members, an increase of 78 academic members.

In order to afford an opportunity to academic members desiring to name additional students as ordinary members of the Association without dues payments by these students, the Board voted to authorize the establishment of two new classes of academic membership, namely:

Contributing Academic Membership, with annual dues of $75 and the privilege of receiving two copies of the MONTHLY and nominating three persons to ordinary membership in the Association (or receiving one copy of the MONTHLY and nominating four persons), and

Sponsoring Academic Membership, with annual dues of $100 and the privilege of receiving two copies of the MONTHLY and nominating six persons to ordinary membership in the Association (or receiving one copy of the MONTHLY and nominating seven persons).

A Sponsoring Academic Member shall have the privilege of naming, for each additional contribution of $25, an additional three persons as ordinary members.

An academic member of any of these classes shall have the privilege of receiving a copy of the MATHEMATICS MAGAZINE in lieu of either one of the copies of the MONTHLY or nominating one person to ordinary membership.

The Board approved a resolution that the Association not sponsor or co-sponsor symposia or conferences designated as honoring an individual mathematician.

ANNUAL BUSINESS MEETING OF THE ASSOCIATION

The Annual Business Meeting was held on Sunday, January 24, 1971, in the Pennsylvania Room of Haddon Hall with President Young presiding. The Association's Tenth Award for Distinguished Service to Mathematics was made to Professor B. W. Jones of the University of Colorado. The citation (which appears on pages 111–112 of the February issue of this MONTHLY) was prepared and read by Professor E. G. Begle of Stanford University. The Award was presented by President Young. Professor Jones, in accepting the Award, stated that he had been flabbergasted when he first heard that this Award was to be made to him, since he felt, in his sane moments, that he was not in the class of those who had previously received it, but, in his less sane moments, he thought of the words of the old song "I do not believe it, but say it again." A specially bound copy of the citation was presented to Mrs. Jones.

The 1971 Chauvenet Prize was presented to Professor Norman Levinson of the Massachusetts Institute of Technology for his paper "A Motivated Account of an Elementary Proof of the Prime Number Theorem," published in this MONTHLY 76 (1969), 225–245. Further details concerning this Prize and its recipient appear on pages 112–113 of the February issue of this MONTHLY. In accepting, Professor Levinson stated that he felt very honored to be selected by the Association for this Prize and could only take refuge in the words of Professor Jones, who had said it very well. He added that he had a bad conscience receiving this Prize since he loved analytic number theory very much, and this paper had been a labor of love. To get a reward for doing it seemed a little too much, but he expressed his appreciation.

The Secretary reported the results of the balloting for Governors of the Association, in which 1851 votes were cast: Professor I. N. Herstein of the University of Chicago and Professor George Springer of Indiana University were elected Governors for the three-year term 1971–73.

The Secretary reported on some of the work undertaken by the Association during the past few months. He expressed the Association's gratitude to the National Science Foundation for having made a grant of $54,250 for a Summer Seminar on the Theory of Probability and Mathematical Statistics to be held at Williams College in the summer of 1971 and for $44,196 in additional support of the Committee on the Undergraduate Program in Mathematics.

He announced an improvement in the Association's financial position as a result of numerous measures adopted in recent months, but, in particular, due to the fact that the Association had received $6,326.80 as contributions and endowment gifts during 1970. In that year, there were 87 Contributing Members, 8 Sponsors, and 5 Patrons. In addition, the MAA has now a total of 47 Life Members, 25 of whom became Life Members in 1970. Six of these Life Members are Patron Life Members.

The Secretary expressed the Association's gratitude to the members of the local Committee on Arrangements under the chairmanship of Professor Albert Schild of Temple University for all their helpful efforts in arranging for this meeting.

The Secretary then moved to amend the By-Laws by deleting from Article IX, Section 2, the words "filed in the office of the Secretary of the State of Illinois and". The motion was approved without dissent.

Professor J. E. Kimber, Jr. presented a resolution concerning the war in Vietnam. President Young requested that the Secretary read that part of the By-Laws describing the purposes of the Association. He then ruled the motion out of order. Upon appeal, this ruling of the chair was sustained by voice vote.

MEETINGS OF OTHER ORGANIZATIONS

The American Mathematical Society held sessions from Thursday, January 21, to Sunday, January 24. The forty-fourth Josiah Willard Gibbs Lecture was delivered by Professor Eberhard Hopf of Indiana University on "Ergodic Theory and the Geodesics on Surfaces of Negative Curvature" on Thursday at 8:30 P.M. Professor Oscar Zariski of Harvard University gave the Retiring Presidential Address on "Some Open Questions in the Theory of Singularities" on Friday at 11:00 A.M. Invited addresses were given by Professor D. P. Sullivan of the Massachusetts Institute of Technology on Thursday at 11:00 A.M. on "Symmetry in Manifold Theory," by Professor D. G. Quillen of the Massachusetts Institute of Technology on Thursday at 1:30 P.M. on "Cohomology of Groups and Algebraic K-Theory", by Professor Leopoldo Nachbin of the University of Rochester and the Instituto de Matematica Pura e Aplicada, Rio de Janeiro, on Friday at 8:30 P.M. on "Recent Developments in Infinite Dimensional Holomorphy", and by Professor Harry Kesten of Cornell University on Sunday at 1:30 P.M. on "Some Non-linear Stochastic Growth Models". There was a special lecture by Dr. Y. L. Luke of the Midwest Research Institute in Kansas City on Sunday at 5:00 P.M. on "Information

Retrieval Systems for Mathematical Journals". All of these lectures were held in the Pennsylvania Room.

The Oswald Veblen Prize in Geometry was awarded on Saturday at 1:15 P.M. in the Pennsylvania Room to Professor R. C. Kirby of the University of California, Los Angeles, for his paper "Stable Homeomorphisms and the Annulus Conjecture", *Ann. Math.* (2), 89(1969), 575–582 and to Professor D. P. Sullivan of the Massachusetts Institute of Technology for his work on the Hauptvermutung summarized in the paper "On the Hauptvermutung for Manifolds", *Bull. Amer. Math. Soc.* 73(1967), 598–600.

The Association for Symbolic Logic met on Thursday and Friday. The Survey Lecture was given by Professor Simon Kochen of Princeton University on Friday at 1:30 P.M. on "Applications of Model Theory to Algebra" in the Pennsylvania Room of Haddon Hall. Invited addresses were given by Professor Kenneth Kunen of the University of Wisconsin, Madison, on Thursday at 3:00 P.M. on "Saturated Ideals" and by Professor Harvey Friedman of the University of Wisconsin, Madison, on Friday at 9:00 A.M. on "Nonstandard Models and Their Applications", all in the Vernon Room of Haddon Hall.

The Conference Board of the Mathematical Sciences sponsored a panel discussion on Operations Research and Mathematics on Saturday at 3:30 P.M. in the Vernon Room of Haddon Hall. This session was planned under the direction of Professor D. L. Iglehart of the Department of Operations Research at Stanford University. Participants in the panel were Professor G. B. Dantzig of the Department of Computer Science and the Department of Operations Research at Stanford University, Professor Julian Keilson of the Department of Statistics at the University of Rochester, Professor T. L. Saaty of the Department of Statistics and Operations Research at the University of Pennsylvania, and Dr. Philip Wolfe of the Mathematical Sciences Department at the Thomas J. Watson Research Center of the International Business Machines Corporation. Professor Iglehart was moderator.

ARRANGEMENTS, ENTERTAINMENT, AND RECREATION

The Committee on Arrangements consisted of Albert Schild, Chairman; H. L. Alder, F. E. Clark, J. P. Clay, Leonard Gillman, W. H. Gottschalk, Sam Newman, C. W. Sloyer, Jr., G. L. Walker.

Registration headquarters were located in the English Lounge of Haddon Hall on the Lounge and Dining Floor. The Mathematical Sciences Employment Register was maintained in the Carolina Room of the Chalfonte Hotel from 9:00 A.M. to 5:00 P.M. on Friday through Sunday, and book and educational media exhibits were displayed in the Exhibit Hall from 9:00 A.M. to 5:00 P.M. on Friday through Sunday.

HENRY L. ALDER, *Secretary*

OFFICERS AND COMMITTEES AS OF FEBRUARY 1, 1971

General Offices: 1225 Connecticut Avenue, N.W., Washington, D. C. 20036
Executive Director: A. B. WILLCOX
Executive Director Emeritus: H. M. GEHMAN
Editorial Director: RAOUL HAILPERN

OFFICERS

President, VICTOR KLEE, University of Washington (1971–72)
Past-President, G. S. YOUNG, University of Rochester (1971)
First Vice-President, GARRETT BIRKHOFF, Harvard University (1970–71)
Second Vice-President, J. W. JEWETT, Oklahoma State University (1971–72)

Editor, HARLEY FLANDERS, Tel Aviv University (1968–73)
Secretary, H. L. ALDER, University of California, Davis (1970–74)
Treasurer, E. A. CAMERON, University of North Carolina (1968–72)

ADDITIONAL MEMBERS OF THE BOARD OF GOVERNORS

Ex-Presidents

E. E. MOISE, Harvard University (1969–74)
R. L. WILDER, University of Michigan (1967–72)
G. S. YOUNG, University of Rochester (1971–76)

Elected Members of the Finance Committee

G. B. PRICE, University of Kansas (1968–71)
G. S. YOUNG, University of Rochester (1971–73)

Governors-at-Large

R. G. BARTLE, University of Illinois, Urbana (1970–72)
R. C. BUCK, University of Wisconsin, Madison (1969–71)
MARY P. DOLCIANI, City University of New York (1969–71)
D. T. FINKBEINER II, Kenyon College (1970–72)
I. N. HERSTEIN, University of Chicago (1971–73)
GEORGE SPRINGER, Indiana University (1971–73)

Editor of the MATHEMATICS MAGAZINE

G. N. WOLLAN, Purdue University (1971–73)

Sectional Governors (July 1, 1968–June 30, 1971)

Florida, D. B. GOODNER, Florida State University
Illinois, F. E. HOHN, University of Illinois
Iowa, D. E. SANDERSON, Iowa State University
Louisiana-Mississippi, S. R. KNOX, Millsaps College
Maryland-D.C.-Virginia, E. E. FLOYD, University of Virginia
Michigan, L. M. KELLY, Michigan State University
North Central, F. L. WOLF, Carleton College
Philadelphia, B. H. BISSINGER, Pennsylvania State University, Middletown
Southern California, R. B. HERRERA, University of California, Los Angeles
Texas, D. E. EDMONDSON, University of Texas, Austin

Sectional Governors (July 1, 1969–June 30, 1972)

Allegheny Mountain, H. L. KRALL, Pennsylvania State University
Indiana, B. E. RHOADES, Indiana University
Kentucky, J. H. WELLS, University of Kentucky
Metropolitan New York, GERALD FREILICH, City University of New York
Nebraska, W. E. MIENTKA, University of Nebraska
Northern California, L. H. LANGE, San Jose State College
Oklahoma-Arkansas, J. E. SCROGGS, University of Arkansas
Rocky Mountain, W. N. SMITH, University of Wyoming
Wisconsin, E. W. SWOKOWSKI, Marquette University

Sectional Governors (July 1, 1970–June 30, 1973)

Kansas, A. M. WEDEL, Ottawa University, Ottawa, Kansas

Missouri, A. G. HADDOCK, University of Missouri, Rolla
New Jersey, P. C. CLIFFORD, Montclair State College
Northeastern, D. E. CHRISTIE, Bowdoin College
Ohio, H. D. LIPSICH, University of Cincinnati
Pacific Northwest, D. W. BUSHAW, Washington State University
Southeastern, HENRY SHARP, JR., Emory University
Southwestern, E. A. WALKER, New Mexico State University
Upper New York State, F. R. OLSON, SUNY College at Fredonia

COMMITTEES OF THE ASSOCIATION

Terms of members expire, except where otherwise noted, at the Annual Meeting in January following the last year of service listed below. For temporary committees, no terms are listed since they are automatically discharged at the expiration of the President's term of office, which is the Annual Meeting in January 1973.

EXECUTIVE COMMITTEE

VICTOR KLEE, *Chairman* (1971–72); H. L. ALDER (1970–74), GARRETT BIRKHOFF (1970–71), E. A. CAMERON (1968–72), HARLEY FLANDERS (1968–73), J. W. JEWETT (1971–72), G. S. YOUNG (1968–71), all *ex officio*.

FINANCE COMMITTEE

VICTOR KLEE, *Chairman* (1971–72), *ex officio;* G. B. PRICE (1968–71), G. S. YOUNG (1971–73), H. L. ALDER (1970–74), *ex officio*, E. A. CAMERON (1968–72) *ex officio*.

COMMITTEE ON ADVISEMENT AND PERSONNEL

K. J. ARNOLD, *Chairman* (1969–71); E. A. DAVIS (1970–72), J. C. EAVES (1969–71), MRS. F. J. MACWILLIAMS (1970–72), G. B. PEDRICK (1971–73), W. H. SCHMIDT (1969–71), W. R. TALBOT (1970–72).

COMMITTEE ON ASSISTANCE TO DEVELOPING COLLEGES

F. M. STEWART, *Chairman* (1970–72); E. M. CARROLL (1970–73), D. E. CHRISTIE (1970–73), L. L. CLARKSON (1970–72), GLORIA C. HEWITT (1970–73), B. J. MARTIN (1970–73), C. O. OAKLEY (1970–71), HENRY SHARP, JR. (1970–73), GEORGE SPRINGER (1970–71), I. E. VANCE (1970–73), K. W. WEGNER (1970–73).

COMMITTEE ON ASSISTANCE TO SECTIONS ON TWO-YEAR COLLEGE PROBLEMS

G. K. GOFF, *Chairman* (1969–72); NORMAN BARTON (1969–72).

COMMITTEE ON CORPORATE MEMBERS

R. E. GASKELL, *Chairman* (1970–74); F. A. BROOKS, JR. (1970–74), JIM DOUGLAS (1970–73), R. L. GRAHAM (1970–73), M. S. KLAMKIN (1970–74), H. E. PICKETT (1970–72), ANDREAS THUSWALDNER (1970–72), LEONARD TORNHEIM (1970–72), A. B. WILLCOX (1970–73).

COMMITTEE ON EARLE RAYMOND HEDRICK LECTURES

E. A. BISHOP, *Chairman* (1970–72); R. D. ANDERSON (1971–73), HARLEY FLANDERS (1969–71).

COMMITTEE ON EDUCATIONAL MEDIA

SEYMOUR SCHUSTER, *Chairman* (1971–73); A. A. BLANK (1969–71), A. G. FADELL (1970–72), B. R. GELBAUM (1969–71), ROBERT KALIN (1969–71), P. E. MILES (1971–73), C. O. OAKLEY (1970–72), H. S. WILF (1970–72), W. B. WOOLF (1970–72).

Committee on High School Contests

J. M. Earl, *Chairman* (1969–71); H. M. Cox, *Director* (1970–73), F. A. Brooks, Jr. (1970–72), L. C. Dalton (1969–71), S. L. Greitzer (1969–71), Alfred Kalfus (1969–72), N. S. Mendelsohn (1969–71), Hans Sagan (1971–73), E. E. Strock (1971–73).

Committee on Institutes

V. O. McBrien, *Chairman* (1970–72); Roy Dubisch (1970–72), D. B. Goodner (1969–71), D. K. Harrison (1969–71), Mrs. Gloria C. Hewitt (1970–72), M. W. Pownall (1969–71), D. L. Thomsen, Jr. (1970–72).

Committee on MAA's Relations with the AAAS

M. M. Day, *Chairman;* R. D. Anderson, F. A. Ficken, J. W. Givens, Jr.

Committee on Publications

Ivan Niven, *Chairman* (1969–71); T. M. Apostol (1970–72), R. G. Bartle (1969–71), E. F. Beckenbach (1971–73), Dorothy L. Bernstein (1969–71), R. P. Boas (1970–72), D. T. Finkbeiner II (1969–71), L. H. Lange (1970–72), Alex Rosenberg (1971–73), E. A. Cameron (1968–72), *ex officio*, Harley Flanders (1968–72), *ex officio*, Joseph Hashisaki (1970–72), *ex officio*, G. N. Wollan (1971–73), *ex officio.*

Subcommittee on Carus Monographs: R. G. Bartle, *Chairman* (1969–71); R. P. Boas (1970–72), D. T. Finkbeiner II (1969–71).

Subcommittee on Lester R. Ford Awards: Ivan Niven, *Chairman* (1969–71) *ex officio;* Marvin Marcus (1971–73), D. E. Richmond (1969–71).

Subcommittee on MAA Studies in Mathematics: Alex Rosenberg, *Chairman* (1971–73); E. F. Beckenbach (1971–73), L. H. Lange (1970–72).

Subcommittee on Slaught Papers: T. M. Apostol, *Chairman* (1970–72); Dorothy L. Bernstein (1969–71), Harley Flanders (1968–73), *ex officio.*

Subcommittee on Miscellaneous Publications: Ivan Niven, *Chairman* (1969–71); H. L. Alder, (1970–74), E. A. Cameron (1968–72), all *ex officio.*

Committee on Publicity

J. W. Brace, *Chairman* (1970–71); L. W. Cohen (1970–71), R. L. Davis (1970–72), G. B. Pedrick (1970–72), G. B. Price (1971–73), F. J. Weyl (1971–73).

Committee on Secondary School Lecturers

H. M. Bacon, *Chairman* (1970–72); Grace E. Bates (1969–71), J. N. Eastham (1970–72), J. R. Hanna (1970–72), J. G. Harvey (1971–73), L. J. Montzingo, Jr. (1971–73), A. D. Stewart (1971–73).

Committee on Sections

L. E. Mehlenbacher, *Chairman* (1971–74); Alfred Aeppli (1971–74), D. W. Blakeslee (1969–72), S. A. Jennings (1968–71), Arnold Wendt (1970–73), A. B. Willcox, *ex officio.*

Committee on Special Funds of the Association

C. B. Allendoerfer, *Chairman* (1969–71); E. A. Cameron (1970–72), C. O. Oakley (1971–73).

Committee on the Award for Distinguished Service to Mathematics

L. W. Cohen, *Chairman* (1970–71); B. W. Jones (1971–73), L. J. Paige (1970–72).

Committee on the Chauvenet Prize

N. D. Kazarinoff, *Chairman* (1969–71); S. S. Chern (1970–72), Norman Levinson (1971–73).

Committee on National Awards and Public Representation

A. W. Tucker, *Chairman* (1969–71) R. P. Boas (1971–73), C. W. Curtis (1971–73), Victor Klee (1971–72) *ex officio;* R. L. Wilder (1971–73).

Committee on the Preparation of a 50-year History of the Association

K. O. May, *Chairman;* C. B. Boyer, D. J. Struik

Committee on the Putnam Prize Competition

W. S. Loud, *Chairman* (1969–71); J. H. McKay, *Director* (1968–72), M. S. Klamkin (1970–72), N. S. Mendelsohn (1971–73).

Committee on the Role of the Two-Year College Mathematics Teachers in the Association

J. W. Jewett, *Chairman;* D. B. Goodner, S. A. Jennings, H. J. Osner, B. E. Rhoades, William Swyter.

Committee on the Undergraduate Program in Mathematics

Alex Rosenberg, *Chairman* (1969–72); Grace E. Bates (1970–72), R. P. Boas (1971–73), D. W. Bushaw (1969–71), L. L. Clarkson (1970–72), E. A. Coddington (1970–72), R. L. Graham (1971–73), I. N. Herstein (1970–72), J. W. Jewett (1970–72), J. L. Kelley (1971–73), D. L. Kreider (1969–71), W. J. LeVeque (1970–72), W. C. Rheinboldt (1970–72), A. E. Ross (1971–73), Dorothy Stone (1969–71), M. D. Thompson (1971–73), J. H. Wells (1971), June P. Wood (1971–73), E. G. Begle, *ex efficio,* Victor Klee (1971–72) *ex officio.*

Advisory Group on Communications: D. T. Finkbeiner II, *Chairman* (1967–72); Meyer Jerison (1971–73), R. H. McDowell (1967–72), G. B. Pedrick (1971–73), J. H. Wells (1970–72).

Panel on Applied Mathematics: M. D. Thompson, *Chairman* (1970–72); F. G. Brauer (1970–71), E. A. Coddington (1970–72), R. L. Graham (1970–71), L. N. Howard (1970–71), D. J. Kleitman (1970–71), D. A. Sanchez (1970–72), H. S. Wilf (1970–71).

Panel on College Teacher Preparation: D. W. Bushaw, *Chairman* (1969–71); Gerald Freilich (1971–73), W. J. LeVeque (1970–72), M. W. Pownall (1969–71), D. R. Sherbert (1971–73), Dorothy Stone (1969–71).

Panel on Computing: W. C. Rheinbolt, *Chairman* (1967–72); R. H. Balomenos (1970–71), S. D. Conte (1970–72), F. E. Hohn (1970–72), T. E. Hull (1970–72), M. E. Rose (1970–71).

Panel on Mathematics in Two-Year Colleges: J. W. Jewett, *Chairman* (1969–72); Grace E. Bates (1970–72), L. J. Fibel (1969–71), R. D. Larsson (1969–71), T. L. McKillips (1971–72), June P. Wood (1970–72), William Wooten (1969–71).

Panel on Special Problems of Minority Groups: L. L. Clarkson, *Chairman* (1970–72); S. H. Douglas (1971), I. N. Herstein (1970–72), F. A. James (1970–72), R. K. Lashoff (1970–71), R. H. Lee (1970–71), G. S. Young (1970–71).

Panel on Statistics: F. A. Graybill, *Chairman* (1968–71); R. A. Bradley (1968–71), Herman Chernoff (1968–71), P. C. Clifford (1968–71), Samuel Goldberg (1968–71), John Neter (1968–71), G. E. Nicholson (1968–71).

Panel on Teacher Training: D. L. Kreider, *Chairman* (1968–71); Shirley A. Hill (1968–71), P. J. Hilton (1968–71), E. F. Krause (1969–71), R. H. McDowell (1969–71), M. E. Shanks (1968–71), R. J. Troyer (1970–72), E. G. Begle, *ex officio.*

Committee on Visiting Lecturers

M. W. Pownall, *Chairman* (1971–73); E. M. Beesley (1969–71), Brindell Horelick (1971–73), W. K. Moore (1969–71), W. M. Myers, Jr. (1969–72), J. R. Wesson (1971–73), W. L. Williams (1969–71).

Committee to Consider Certification and Accreditation in Mathematics

D. T. Finkbeiner II, *Chairman;* R. G. Bartle, Grace E. Bates, E. A. Cameron, Joseph Hashisaki, L. H. Lange, F. L. Wolf.

Film Projects Advisory Committee and Advisory Committee for Individual Lecture Films

A. A. Blank, *Chairman* (1970–72); R. G. Long, *Director* (1970–72), E. F. Beckenbach (1970–72), R. C. Buck (1971–73), L. W. Cohen (1971–73), R. C. Fisher (1970–72), R. D. James (1970–72), P. D. Lax (1971–73), B. J. Pettis (1970–72), Seymour Schuster (1971–73), H. N. Shapiro (1971–73).

Joint Committee on Employment Opportunities

Terms of members of this committee expire on February 28 of the last year of service listed.

R. J. Thompson, *Chairman* (1968–72, MAA), J. I. Rosenblatt (1970–72, AMS), R. A. Handelsman (1971–72, SIAM).

Joint Committee on Places of Meetings

G. L. Walker, *Chairman;* H. L. Alder, Everett Pitcher, A. B. Willcox, all *ex officio.*

Nominating Committee for 1971

C. B. Allendoerfer, *Chairman;* R. C. Buck, D. B. Goodner.

EDITORIAL BOARDS OF THE ASSOCIATION

American Mathematical Monthly (all terms expire December 31, 1973).
Editor: Harley Flanders.
Associate Editors: Joshua Barlaz, E. R. Berlekamp, David Drasin, R. K. Guy, Raoul Hailpern, J. G. Harvey, P. D. Lax, A. P. Mattuck, G. P. Murphy, M. W. Pownall, Gian-Carlo Rota, Seymour Schuster, J. A. Seebach, Jr., E. P. Starke, L. A. Steen, J. G. Wendel.

Mathematics Magazine (all terms expire December 31, 1973).
Editor: G. N. Wollan
Associate Editors: L. C. Eggan, Howard Eves, Raoul Hailpern, R. E. Horton, D. Elizabeth Kennedy, Hans Sagan, B. L. Schwartz, P. J. Zwier.

REPRESENTATIVES OF THE ASSOCIATION

On the Advisory Committee to the Topology Films Project:
G. S. Young (1970–72)
On the AAAS Cooperative Committee on the Teaching of Mathematics and Science:
M. W. Pownall (1969–71)
On the American Council on Education:
H. L. Alder, *ex officio*, Victor Klee, *ex officio*
On the Conference Board of the Mathematical Sciences:
H. L. Alder, *ex officio*, Victor Klee, *ex officio*
On the Council of the American Association for the Advancement of Science:
J. W. Brace (1971–73), M. M. Day (1969–71)

On the Editorial Board of the TYC MATHEMATICS JOURNAL:
JOSEPH HASHISAKI (1970–72), J. W. JEWETT (1969–72)
On the Governing Council of Mu Alpha Theta:
G. B. PRICE (1970–72)
On the National Research Council:
E. J. MCSHANE (July 1, 1968–June 30, 1971)
On the U. S. Commission on Mathematical Instruction:
A. A. BLANK (July 1, 1969–June 30, 1973), P. J. HILTON (July 1, 1970–June 30, 1974).

ACADEMIC MEMBERS ELECTED INTO THE ASSOCIATION

In accordance with the amendment adopted at the business meeting of the Association at Stillwater, Oklahoma, on August 30, 1961, the Board of Governors at its meeting in Atlantic City, New Jersey, on January 22, 1971, elected to membership the nineteenth set of applicants for academic membership (for election of the other eighteen sets, see the March and December issues of 1969 and the April and November issues of 1970). Approval for election was given to the following twenty-six applicants for academic membership:

Atlantic Community College, Mays Landing, New Jersey
Boston College, Boston, Maine
Cambrian College, Sudbury, Ontario, Canada
Carleton College, Ottawa, Ontario, Canada
Community College of Allegheny, West Mifflin, Pennsylvania
Delaware State College, Dover, Delaware
Eisenhower College, Seneca Falls, New York
Genesee Community College, Batavia, New York
George Brown College of Applied Arts & Technology, Toronto, Ontario, Canada
Goucher College, Baltimore, Maryland
Housatonic Community College, Stratford, Connecticut
LeMoyne College, Syracuse, New York
McMaster University, Hamilton, Ontario, Canada
Middlesex Community College, Middletown, Connecticut
Northeastern University, Boston, Massachusetts
Temple Buell College, Denver, Colorado
University of Alabama, Huntsville, Alabama
University of Arkansas, Little Rock, Arkansas
University of Guelph, Guelph, Ontario, Canada
University of Lethbridge, Lethbridge, Alberta, Canada
University of Tennessee, Chattanooga, Tennessee
University of Waterloo, Waterloo, Ontario, Canada
Upsala College, East Orange, New Jersey
Virginia Commonwealth University, Richmond, Virginia
Wichita State University, Witchita, Kansas
York University, Downsview, Ontario, Canada

HENRY L. ALDER, *Secretary*

NOVEMBER MEETING OF THE NEW JERSEY SECTION

The Fall meeting of the New Jersey Section of the MAA was held at Rutgers, The State University, on November 14, 1970. Professor Hal Trotter, Princeton University, Chairman of the Section, presided at the morning meeting. Seventy persons attended the meeting, including forty-five members of the MAA.

The following papers were presented during the morning session:

Generalized quantifiers, by Michael Aissen, Rutgers University.

Circles and fixed points, by Michael Atiyah, Institute for Advanced Study.

The afternoon session was chaired by S. L. Greitzer of Rutgers, The State University, and consisted of a panel meeting on "After Cambridge—What?" The panel members were: H. F. Fehr, Teachers College, Columbia University; Harry Ruderman, Hunter College H. S.; Franklin Armour, Thomas Jefferson J. H. S., Teaneck, N. J.; Richard Krutch, Hunter College H. S.

In a short business meeting, Samuel Greitzer, Rutgers, The State University, was elected Chairman, and F. J. Almgren, Princeton University, was elected to the council.

JOHN RECKZEH, *Secretary-Treasurer*

MATHEMATICAL SCIENCES EMPLOYMENT REGISTER

Applicant qualification forms and position description forms are now available for those persons who wish to list in the May issue of the Mathematical Sciences Employment Register. The deadline for receipt of the completed forms is April 1, 1971. There is no charge for listing in the published list except when the late listing charge of $5 is applicable. Provision may be made for anonymity of applicants upon payment of $5 to defray the cost involved in handling such a listing; this fee must be submitted with the applicant qualification form. Forms may be obtained from the Mathematical Sciences Employment Register, Post Office Box 6248, Providence, Rhode Island 02904.

An open Register will be maintained at the meeting of the American Mathematical Society to be held on April 7–10, 1971. The Register will be open from 9:00 A.M. to 5:00 P.M. on Thursday and Friday, April 8–9, in the Jade Room of the Waldorf-Astoria in New York City. A complete announcement of this meeting appears in the February issue of the NOTICES. As this is not a joint meeting with the other sponsoring organizations, the costs of this Register will be borne entirely by the Society. To assist the Register personnel in planning for this open Register, it is requested that those persons who expect to participate send a note to the Providence office. Employers are requested to indicate the number of positions available and to specify which day or days they will be available for interviewing.

The May issue of the Register will be mailed from Providence in late April. A subscription to the lists, which includes three issues (May, August, and January) of both the applicants list and the positions list, is available for $30 a year; the individual issues of both lists may be purchased in May, August, and January for $15. A subscription to the applicants list alone or single copies of that list are not available. Copies of the positions list only may be purchased for $5. A subscription to the list of positions, which also includes three issues (May, August, and January), is available for $12 a year.

It should be noted that the lists are mailed "Book Rate" (average delivery time from Providence to most locations is approximately 14 to 21 days or longer) unless the purchaser either indicates a willingness in advance to pay the "First Class" or "Air Mail" charges, or includes the fee for this service when prepayment is made. The applicable postage charges, determined by the location of the purchaser, will be furnished on request to those persons who would like to take advantage of this service.

NEW EDITION OF PROFESSIONAL OPPORTUNITIES IN MATHEMATICS

The eighth edition of PROFESSIONAL OPPORTUNITIES IN MATHEMATICS was published in February, 1971. It is a completely rewritten and updated version of a publication which has been in continuous existence since 1951. Its 27 pages contain discussions of The Mathematician in Industry, The Mathematician in Government, Opportunities in Applied Probability and Mathematical Statistics, Opportunities in the Actuarial Profession, and Job Opportunities for persons who have completed the Undergraduate Mathematics Major. There is a bibliography containing 33 references for further reading on careers in Mathematics.

PROFESSIONAL OPPORTUNITIES IN MATHEMATICS may be purchased for 35¢ (30¢ in lots of 5 or more) from the Association's Washington Office.

CALENDAR OF FUTURE MEETINGS

Fifty-second Summer Meeting, Pennsylvania State University, University Park, August 30-September 1, 1971.

Fifty-fifth Annual Meeting, Las Vegas, Nevada, January 19–21, 1972.

The following is a list of the Sections of the Association with dates of future meetings so far as they have been reported to the Editorial Director.

ALLEGHENY MOUNTAIN, Geneva College, Beaver Falls, Pennsylvania, May 7–8, 1971.

FLORIDA

ILLINOIS, Eastern Illinois University, Charleston, May 14–15, 1971.

INDIANA, Purdue University, North Central Campus, Westville, May 8, 1971.

IOWA

KANSAS

KENTUCKY

LOUISIANA-MISSISSIPPI

MARYLAND-DISTRICT OF COLUMBIA-VIRGINIA

METROPOLITAN NEW YORK

MICHIGAN, Western Michigan University, Kalamazoo, May 7–8, 1971.

MISSOURI

NEBRASKA

NEW JERSEY

NORTH CENTRAL, University of Minnesota, Minneapolis, May 8, 1971.

NORTHEASTERN, Colby College, Waterville, Maine, June 19, 1971.

NORTHERN CALIFORNIA

OHIO

OKLAHOMA-ARKANSAS

PACIFIC NORTHWEST, Oregon State University, Corvallis, June 18–19, 1971.

PHILADELPHIA, Lafayette College, Easton, November 20, 1971.

ROCKY MOUNTAIN, Weber State College, Ogden, Utah, May 7–8, 1971.

SOUTHEASTERN

SOUTHERN CALIFORNIA

SOUTHWESTERN

TEXAS

UPPER NEW YORK STATE, St. Lawrence University, Canton, May 8, 1971.

WISCONSIN

FUTURE MEETINGS OF OTHER ORGANIZATIONS

AMERICAN ASSOCIATION FOR THE ADVANCEMENT OF SCIENCE, Philadelphia, December 26–31, 1971.

AMERICAN MATHEMATICAL SOCIETY, Pennsylvania State University, University Park, August 31–September 3, 1971.

AMERICAN SOCIETY FOR ENGINEERING EDUCATION, U. S. Naval Academy, Annapolis, June 21–24, 1971.

ASSOCIATION FOR COMPUTING MACHINERY, Chicago, August 3–5, 1971.

ASSOCIATION FOR SYMBOLIC LOGIC

CENTRAL ASSOCIATION OF SCIENCE AND MATHEMATICS TEACHERS, Detroit, Michigan, November 25–27, 1971.

FIBONACCI ASSOCIATION, College of the Holy Names, Oakland, California, November 13, 1971.

INSTITUTE OF MATHEMATICAL STATISTICS, Fort Collins, Colorado, August 23–26, 1971.

MU ALPHA THETA, Pennsylvania State University, University Park, September 1, 1971.

NATIONAL COUNCIL OF TEACHERS OF MATHEMATICS, Chicago, Illinois, April 16–20, 1972.

OPERATIONS RESEARCH SOCIETY OF AMERICA, Sheraton Dallas, Dallas, May 5-7, 1971.

PI MU EPSILON, Pennsylvania State University, University Park, August 31–September 1, 1971.

SOCIETY FOR INDUSTRIAL AND APPLIED MATHEMATICS, Seattle, Washington, June 28–30, 1971.

GEORGE BANTA COMPANY, INC., MENASHA WISCONSIN

THE AMERICAN MATHEMATICAL MONTHLY

(Founded in 1894 by Benjamin F. Finkel)

THE OFFICIAL JOURNAL OF

THE MATHEMATICAL ASSOCIATION OF AMERICA, INC.

VOLUME 78 NUMBER 5

CONTENTS

(Continued on inside cover)

MAY 1971

NOTICE TO AUTHORS

Specialized research is usually unsuitable; see *Statement of Policy* (vol. 76, p. 2). Manuscript preparation: Please use the *Manual for Monthly Authors* (vol. 78, p. 1) and follow the format in current issues of the MONTHLY. Manuscripts should be typewritten, triple-spaced with wide margins; submit two copies and keep one for protection against loss. Each author receives 50 free reprints.

Backlog: Main Articles 9 months, Math. Notes 9 months, Research Problems 7 months, Classroom Notes 8 months, Math. Education 6 months.

EDITORIAL CORRESPONDENCE AND MAIN ARTICLES: to HARLEY FLANDERS, American Mathematical Monthly, Tel Aviv University, Ramat Aviv, Israel (see Notice, vol. 77, 1970, p. 555); NOTES, etc.: to the corresponding Associate Editor; ADVERTISING CORRESPONDENCE: to RAOUL HAILPERN, Mathematical Association of America, SUNY at Buffalo, Buffalo, N. Y. 14214; CHANGE OF ADDRESS and SUBSCRIPTIONS: to A. B. WILLCOX, Mathematical Association of America, 1225 Connecticut Ave., N.W., Washington, D. C. 20036

Annual dues for members of the Association (including a subscription to the American Mathematical Monthly) are $12.50. For nonmembers the subscription price is $18.00.

PUBLISHED BY THE ASSOCIATION at Washington, D. C., and Menasha, Wisconsin during the months of January, February, March, April, May, June–July, August–September, October, November, December.

Second-class postage paid at Washington, D. C., and additional mailing offices. Acceptance for mailing at special rate of postage provided for in the Act of February 28, 1925, embodied in Paragraph 4, Section 538, P. L. and R., authorized April 1, 1926.

WHAT IS A MARTINGALE?

J. L. DOOB, University of Illinois

1. Introduction. Martingale theory illustrates the history of mathematical probability: the basic definitions are inspired by crude notions of gambling, but the theory has become a sophisticated tool of modern abstract mathematics, drawing from and contributing to other fields. Martingales have been studied systematically for about thirty years, and the newer probability texts usually devote some space to them, but the applications are so varied that there is no one place where a full account can be found. References [1] and [2] are the most complete sources.

The following account of martingale theory is designed to give a feeling for the subject with a minimum of technicality. The basic definitions are given at two levels, of which the first is more intuitive and elementary and suffices for some of the examples. The examples illustrate only the immediate applications requiring a minimum of background.

We recall that in probability theory one starts with a set called the **sample space,** that **events** are subsets of this space, and **random variables** are functions on this space. Suppose for simplicity that the sample space Ω has only countably many points $\omega_1, \omega_2, \cdots$ to which are assigned probabilities $p_1, p_2, \cdots$ respectively, with $p_j \geqq 0$ and $\sum_j p_j = 1$. If $x_1, \cdots, x_k$ are random variables, we write

$$\{x_1 = a_1, \cdots, x_k = a_k\} = \bigcap_m \{\omega_j : x_m(\omega_j) = a_m\}$$

for the set of points where the random variables have the indicated values. The **probability** $P\{A\}$ of the event A is defined as $\sum p_j$, where the sum is over those values of j with ω_j in A. If $P\{B\} > 0$, the **conditional probability** of the event A relative to B is defined by $P\{A \mid B\} = P\{A \cap B\}/P\{B\}$. If x is a random variable, its **expectation** is defined as

$$E\{x\} = \sum_j x(\omega_j)p_j, \tag{1.1}$$

(where it is supposed that the series converges absolutely) and the **conditional expectation** of x relative to B is defined correspondingly when $P\{B\} > 0$ as

$$E\{x \mid B\} = \sum_j{}' x(\omega_j)p_j/P\{B\}, \tag{1.2}$$

Professor Doob received his Harvard Ph.D. under J. L. Walsh, spent the next two years at Columbia on an N.R.C. Fellowship, a third year there on a Carnegie Corp. grant to Harold Hotelling, and began his long association with the University of Illinois in 1935. He is a member of the National Academy of Sciences, served as President of the American Mathematical Society in 1963–64, and served the Inst. of Math. Stat. as Vice-President in 1945 and as President in 1950. He is the author of *Stochastic Processes* (Wiley, 1953) and of numerous important research papers on probability theory and applications. *Editor.*

where the prime indicates that the sum is over the values of j with ω_j in B. If $P\{B\}=0$ the preceding conditional probability and expectation can be defined arbitrarily without affecting later work.

It has been found useful to make conditional expectations into functions, as follows. Let $\{B_n, n\geqq 1\}$ be a partition of Ω, that is, a countable class of disjoint sets with union Ω. This partition generates, and is in turn determined by, a σ-algebra $\mathfrak{F}$, namely the class of all unions of sets of the partition. If x is a random variable with an expectation, define $E\{x\mid\mathfrak{F}\}$, the **conditional expectation** of x relative to $\mathfrak{F}$, as the random variable with the constant value $E\{x\mid B_n\}$ on each set B_n. This definition is unambiguous except on the partition sets (if there are any) of probability 0. In particular if $y_1, \cdots, y_k$ are random variables, they induce the partition $\mathfrak{F}$ each of whose sets is determined by a condition of the form $\{y_1=a_1, \cdots, y_k=a_k\}$; in this case $E\{x\mid\mathfrak{F}\}$, also denoted by $E\{x\mid y_1, \cdots, y_k\}$, is the function with the value $E\{x\mid y_1=a_1, \cdots, y_k=a_k\}$ on the set $\{y_1=a_1, \cdots, y_k=a_k\}$.

Let $\mathfrak{F}_1\subset\mathfrak{F}_2\subset\cdots$ be a finite or infinite increasing sequence of σ-algebras (generated by partitions of the sample space as just described). The intuitive picture to keep in mind is that $\mathfrak{F}_n$ represents the class of all relevant past events up to and including time n. The monotoneity relation corresponds to the idea that the past to time $n+1$ includes more events than the past to time n. Let $x_1, x_2, \cdots$ be a sequence of random variables. We consider x_n as part of the relevant history to time n, interpreting this statement to mean that each event of the form $\{x_n=a\}$ is a set in the class $\mathfrak{F}_n$. The sequence of random variables is to be analyzed. In some applications $x_1, x_2, \cdots$ are specified and $\mathfrak{F}_n$ is the past as determined entirely by $x_1, \cdots, x_n$, that is $\mathfrak{F}_n$ is generated by the partition of Ω induced by $x_1, \cdots, x_n$. This choice of $\mathfrak{F}_n$ will be called **minimal** (relative to a specified sequence of random variables).

The sequence $\{x_n, n\geqq 1\}$ is called a **martingale** relative to $\{\mathfrak{F}_n, n\geqq 1\}$ if each x_n has an expectation, and if for $m<n$ the expected value of x_n given the past up to time m is x_m, that is

$$E\{x_n\mid\mathfrak{F}_m\} = x_m. \tag{1.3}$$

This is a relation between functions on the sample space and is to hold almost everywhere, that is everywhere except perhaps on a subset of the sample space of probability 0. If every p_j is strictly positive, the exceptional set is empty. If x_n is thought of as the fortune of a gambler at time n, the defining equality (1.3) corresponds to the idea that the game the gambler is playing is fair. If '$=$' in (1.3) is replaced by '$\geqq$' or '$\leqq$', the sequence of random variables is called a **submartingale** or **supermartingale** respectively, and the corresponding games are then respectively favorable or unfavorable to the gambler. Trivially (for specified σ-algebras) $\{x_n, n\geqq 1\}$ is a supermartingale if and only if $\{-x_n, n\geqq 1\}$ is a submartingale, and is a martingale if and only if it is both a supermartingale and a submartingale. The definitions imply that $E\{x_n\}$ increases with n in the submartingale case, decreases with n in the supermartingale case, and does not

vary with n in the martingale case. (All monotoneity statements are to be interpreted in the wide sense.)

Equation (1.3) implies that

$$E\{x_n \mid x_1, \cdots, x_m\} = x_m, \tag{1.4}$$

equivalently that

$$E\{x_n \mid x_1 = a_1, \cdots, x_m = a_m\} = a_m, \tag{1.4'}$$

whenever the conditioning event has strictly positive probability, and in fact (1.4) is the same as (1.3) whenever every $\mathfrak{F}_k$ is minimal. In other words, a martingale relative to a given sequence of σ-algebras is also one relative to the minimal sequence of σ-algebras. A corresponding remark is valid for submartingales and supermartingales. If the sequence of σ-algebras is not mentioned, the minimal sequence is to be understood.

The definition of a martingale is applicable to complex-valued random variables, and we shall consider certain complex martingales below. Trivially, the real and imaginary parts of a complex martingale are real martingales.

2. Definitions in the general case. The definitions in Section 1 assumed countability of the sample space, a condition not satisfied for some of the applications to be described below. In this section definitions will be given in the general case, in non-probabilistic language to convince cynical readers that probability theory does not need an admixture of non-mathematical terms like *coin*, *event*, *gambler*, *urn*, $\cdots$, even though the ideas behind these terms have inspired much of the theory.

Let $\{\Omega, \mathfrak{F}, P\}$ be a measure space: Ω is a set, $\mathfrak{F}$ is a σ-algebra of subsets of Ω, and P is a measure defined on $\mathfrak{F}$. Assume further that $P\{\Omega\}=1$. (Some non-probabilists accuse probabilists of seeking to mystify outsiders by disguising measurable functions and their integrals with the aliases *random variables* and *expectations*. Note however that probabilists were dealing with the integrals of functions on abstract sets before other analysts dreamed of measure theory. It is sardonic that, dually, some probabilists accuse others of obfuscating probability with measure theory.) Let $\mathfrak{F}_1 \subset \mathfrak{F}_2 \subset \cdots$ be an increasing sequence of σ-algebras of $\mathfrak{F}$ sets. Let $\{x_n, n \geqq 1\}$ be a sequence of complex functions on Ω satisfying the following conditions:

(a) x_n is measurable relative to $\mathfrak{F}_n$;
(b) x_n is integrable;
(c) If $m < n$ and if A is any set in $\mathfrak{F}_m$, then

$$\int_A x_n dP = \int_A x_m dP. \tag{2.1}$$

Then the sequence of random variables is said to be a **martingale** relative to $\{\mathfrak{F}_n, n \geqq 1\}$. If the random variables are real and if '=' is replaced in (2.1) by '$\leqq$'

or '$\geqq$', the sequence of functions is said to be a **supermartingale** or **submartingale,** respectively, relative to the sequence of σ-algebras. Conditional expectations relative to a σ-algebra (in the present general context) are defined in such a way that (2.1) and (1.3) (to hold P almost everywhere on Ω) are equivalent. Thus the present definitions include the earlier ones and the collateral remarks about martingale theory in Section 1 are valid in the general case also.

The definitions have been given for the parameter set $1, \cdots, k$ or $1, 2, \cdots$, ordered as usual, but they are obviously extendable to any simply-ordered parameter set.

3. Example: expectations knowing more and more. Let $\mathfrak{F}_1 \subset \mathfrak{F}_2 \subset \cdots$ be an increasing sequence (finite or infinite) of σ-algebras, as usual, and let x be a random variable with an expectation. Then if $x_n = E\{x \mid \mathfrak{F}_n\}$, the sequence $x_1, x_2, \cdots$ is a martingale relative to the given σ-algebra sequence. That is, successive conditional expectations of x, as we know more and more, yield a martingale. More generally, the parameter set $1, 2, \cdots$ can be replaced in this example by any simply ordered set. The essential condition is that $\mathfrak{F}_t$ increase with t. Every martingale whose parameter set has a last element is of this type, with x identified with the last element.

This example suggests the possibility of applications to statistics and information theory and induces such probabilistic extravagances as the statement that 'the game man plays with nature as he learns more and more is fair.' (Perhaps this statement indicates how realistically a martingale reproduces the idea of a fair game.)

4. Example: sums of independent random variables. Let $y_1, y_2, \cdots$ be independent random variables with expectations, and let $x_n = y_1 + \cdots + y_n$. Then it is intuitively obvious and easily proved that $x_1, x_2, \cdots$ is a martingale if $E\{y_j\} = 0$ for $j > 1$, a submartingale if $E\{y_j\} \geqq 0$ for $j > 1$, a supermartingale if $E\{y_j\} \leqq 0$ for $j > 1$.

5. Example: averages of independent random variables. In the preceding example suppose that $y_1, y_2, \cdots$ have a common distribution which has an expectation. Then a symmetry argument shows that

$$\cdots, \frac{x_3}{3}, \frac{x_2}{2}, x_1$$

is a martingale in the indicated order (left to right). This example suggests possible applications to the law of large numbers and therefore suggests ties between martingale theory and ergodic theory. In fact there are close relationships between the two theories, and sometimes it is said that one contains the other. Which is said to contain the other depends on the speaker. At any rate, it is true that in a reasonable sense there are only two qualitative convergence theorems in measure theory (aside from theorems of the form "convergence of type I implies convergence of type II"), the ergodic theorem and the martingale

convergence theorem. The latter will be discussed below. Each is involved with finer and finer averaging.

6. Example: harmonic functions on a lattice. Let S be any subset of the set S' of points with integral coordinates in d-dimensional coordinate space, $d \geqq 1$. A point ξ of S will be called an **interior** point of S if S contains all $2d$ of the nearest neighbors in S' of ξ. Otherwise ξ will be called a **boundary** point of S. Unless $S=S'$ there will be boundary points. A function u on S will be called **harmonic** (**superharmonic**) if u at each interior point of S is equal (at least equal) to the average of u on its $2d$ nearest neighbors. For example, a linear function is harmonic, a concave function of a harmonic function is superharmonic. Define a **walk** on S, that is, a sequence $x_0, x_1, \cdots$ of random variables with values in S, as follows. Prescribe some initial point in S and set x_0 identically this point. If $x_0=a_0, \cdots, x_n=a_n$ and if a_n is an interior point of S, then x_{n+1} is to be (conditional probability) one of the $2d$ nearest neighbors of a_n, with (conditional) probability $1/(2d)$ for each one. If a_n is a boundary point of S, x_{n+1} is to be a_n. Thus the walk proceeds until the boundary is reached, if ever, and sticks at the first boundary point reached. It can be shown that such a walk exists. If u is harmonic (superharmonic) on S, the sequence of random variables $u(x_0)$, $u(x_1), \cdots$ is a martingale (supermartingale). If we are to consider the infinite sequence $x_0, x_1, \cdots$, the sample space must be uncountable.

7. Example: classical harmonic and analytic functions. A variation of the idea of Section 6 is the following. Let S be an open subset of d-dimensional coordinate space, $d \geqq 1$. A function u on S is said to be **harmonic** if u is continuous and if, whenever ξ is a point of S, and B is a ball with center ξ whose closure lies in S, the value of u at ξ is the average of its values on the boundary of B. For example linear functions are harmonic for all d, and are the only harmonic functions if $d=1$ and S is an interval. The harmonic functions are the infinitely differentiable functions whose Laplacians vanish. If $d=2$, the real part of an analytic function is harmonic. If ξ is in S and if S is the whole space, denote by $B(\xi)$ the boundary of the ball with center ξ and radius 1. If S is not the whole space, denote by $B(\xi)$ the boundary of the ball with center ξ and radius half the distance from ξ to the boundary of S. If $d>1$ and $A \subset B(\xi)$, let $p(\xi, A)$ be the $(d-1)$-dimensional "area" of A divided by that of $B(\xi)$. If $d=1$ let $p(\xi, A)$ be one-half the number of points in A. Now define a walk on S, that is, random variables $x_0, x_1, \cdots$ with values in S, as follows. Prescribe some initial point in S and set x_0 identically this point. If $x_0=a_0, \cdots, x_n=a_n$, then x_{n+1} is to be on $B(a_n)$, and in fact the conditional probability that x_{n+1} is in the subset A of $B(a_n)$ is to be $p(a_n, A)$. With this definition, if u is real and harmonic, or if $d=2$ and u is complex and analytic on S, the sequence of random variables $\{u(x_n), n \geqq 0\}$ is a martingale. If superharmonic and subharmonic functions are defined as usual in this context, there is a corresponding relation between superharmonic (subharmonic) functions and supermartingales (submartingales).

Sections 6 and 7 indicate connections between martingale theory and

potential theory. In fact the probability theory of Markov processes and abstract potential theory are to a considerable extent different ways of looking at the same subject, and martingale theory is an essential tool of probabilistic potential theory.

8. Two basic principles. We shall need the concept of a **stopping time,** also called a **Markov time** and **optional time.** If $\mathfrak{F}_0 \subset \mathfrak{F}_1 \subset \cdots$ is an increasing sequence of σ-algebras and if ν is a random variable whose range is the set of positive integers together with $+\infty$, the random variable ν is called a **stopping time** (relative to the sequence of σ-algebras) if $\{\nu \leqq k\}$ is a set in $\mathfrak{F}_k$ for $k=0, 1, \cdots$, that is, in intuitive language, if the condition $\nu \leqq k$ is a condition involving only what has happened up to and including time k. For example, if $\{x_n, n \geqq 0\}$ is a martingale relative to the given sequence of σ-algebras and if $\nu(\omega)$ is the first integer j for which $x_j(\omega) > 0$, or $\nu(\omega) = \infty$ if there is no such integer, then ν is a stopping time.

Two basic principles are embodied in the following rough statements, which will be given exact versions and applied in various contexts. Let $x_0, x_1, \cdots$ be a supermartingale relative to $\mathfrak{F}_0, \mathfrak{F}_1, \cdots$.

P_1. *If $\nu_1 \leqq \nu_2 \leqq \cdots$ are finite stopping times which are not too large (with reference to $x_0, x_1, \cdots$), then the sequence $x_{\nu 1}, x_{\nu 2}, \cdots$ is a supermartingale; it is a martingale if $\{x_n, n \geqq 0\}$ is a martingale.*

P_2. *If $\sup_n E\{|x_n|\}$ is not large, then $\sup_n |x_n|$ is not large, and the sequences $\{x_0(\omega), x_1(\omega), \cdots\}$ of possible values of the supermartingale are not strongly oscillatory.*

The principle P_1 is suggested by the fact that a gambler playing an unfair (fair) game will still consider it unfair (fair) if he looks at his money not after every play, but only after plays number $\nu_1, \nu_2, \cdots$. The following version of P_1, in which all stopping times are finite, will be called the STOPPING TIME THEOREM below:

If $\{x_n, n \geqq 0\}$ is a supermartingale (martingale) and if each ν_j is bounded, then the sequence $\{x_{\nu_j}, j \geqq 1\}$ is also a supermartingale (martingale). With no restriction on the finite stopping times, the second sequence is a supermartingale if the first was, and if also $x_j \geqq 0$ for all j.

The principle P_2 is exemplified by the following theorem, which will be called the MARTINGALE CONVERGENCE THEOREM below:

If $\{x_n, n \geqq 1\}$ is a supermartingale with $\sup_n E\{|x_n|\} < \infty$, then $\lim_{n\to\infty} x_n$ exists and is finite almost surely. If $\{\cdots, x_{-1}, x_0\}$ (ordered left to right) is a supermartingale, then (almost surely) $\lim_{n\to-\infty} = x_{-\infty}$ exists with $-\infty < x_{-\infty} \leqq \infty$.

Here **almost surely** means everywhere on the sample space except possibly for a set of probability zero. (One of the most noticeable distinctions between

probability and measure theory is that a probabilist frequently writes "almost surely" where a measure theorist writes "almost everywhere.")

The Martingale Convergence Theorem should be contrasted with the ERGODIC THEOREM, which we state in the following version:

Let $z_1, z_2, \cdots$ be random variables with expectations and having the property that for every $n \geqq 0$ the joint distribution of $z_j, \cdots, z_{n+j}$ does not depend on $j \geqq 1$. Then almost surely

$$\lim_{n\to\infty} (z_1 + \cdots + z_n)/n \tag{8.1}$$

exists and is finite.

Denote by z_n' the ratio in (8.1). Then z_n' is a weighted average of $z_1, z_2, \cdots$, and z_{n+1}' is a weighted average of $z_1', \cdots, z_n', z_{n+1}, \cdots$. Thus z_{n+1}' is a coarser average than z_n'. Now the defining equality of a martingale $\{x_n, n \geqq 1\}$ makes x_n a partial average of x_{n+1}. Thus if there is any relation between the Martingale Convergence Theorem and the Ergodic Theorem, one would conjecture that the Ergodic Theorem corresponds to the Martingale Convergence Theorem with decreasing index. The application in Section 11 verifies this conjecture in a special case.

9. Continuation of Section 3. In the example in Section 3, $E\{|x_n|\} \leqq E\{|x|\}$. Thus according to the Martingale Convergence Theorem, $\lim_{n\to\infty} E\{x \mid \mathfrak{F}_n\}$ exists almost surely. If $\mathfrak{F}_\infty$ is defined as the smallest σ-algebra containing every set of $\bigcup_n \mathfrak{F}_n$, then the limit can be identified as $E\{x \mid \mathfrak{F}_\infty\}$. We have obtained a kind of continuity theorem for conditional expectations. There is a corresponding theorem relative to a decreasing sequence of σ-algebras.

10. Continuation of Section 4. Suppose in Section 4 that $E\{y_j\} = 0$ for all j. Then the sequence $\{x_n, n \geqq 1\}$ is a martingale, and according to the Martingale Convergence Theorem, the series $\sum_j y_j$ converges almost surely if $\sup_n E\{|\sum_1^n y_j|\} < \infty$. Suppose for example that $E\{y_j^2\} < \infty$ for all j, so that the series $\sum_j y_j$ is a series of orthogonal random variables, and as such converges in the mean if and only if $\sum_j E\{y_j^2\} < \infty$, that is, if and only if $\sup_n E\{x_n^2\} < \infty$. But the finiteness of this supremum means that the hypothesis of the Martingale Convergence Theorem is satisfied, and we have proved that convergence in mean of a sum of independent random variables with zero expectations implies almost sure convergence.

To obtain a second application of martingale theory to sums of independent random variables, let $z_1, z_2, \cdots$ be mutually independent random variables with a common distribution having expectation α. If $x_0 = 0$ and $x_n = \sum_1^n (z_j - \alpha)$ for $n > 0$, the sequence $\{x_n, n \geqq 0\}$ is a sequence of sums of independent random variables with zero expectations and as such is a martingale. If ν is a finite stopping time (relative to the minimal σ-algebra sequence) with an expectation,

and if $\nu_1=0$, $\nu_2=\nu$, it can be shown that a version of P_1 yields the fact that x_{ν_1}, x_{ν_2} is a martingale with two random variables, having common expectation 0. The fact that $E\{x_\nu\}=0$ means that

$$E\left\{\sum_1^\nu z_j\right\} = \alpha E\{\nu\}. \tag{10.1}$$

This equality is Wald's Fundamental Theorem of Sequential Analysis, which has many applications in statistics.

11. Continuation of Section 5. An application of the Martingale Convergence Theorem to the martingale in Section 5 yields the almost sure existence of the limit

$$\lim_{n\to\infty} (y_1 + \cdots + y_n)/n. \tag{11.1}$$

The limit can be shown to be $E\{y_1\}$. This convergence result is known to probabilists as the Strong Law of Large Numbers for Independent Random Variables with a Common Distribution, and the result can also be obtained as an application of the Ergodic Theorem. (See the discussion of the relation between the Ergodic Theorem and the Martingale Convergence Theorem in Section 8.)

12. Continuation of Section 6. We suppose that S is bounded and show first that almost every walk path from a point ξ of S reaches the boundary. There are elementary proofs of this fact and the following proof is given only to illustrate martingale theory. If u is defined and harmonic on S, then $\{u(x_n),\ n\geqq 0\}$ is a bounded martingale and is therefore almost surely convergent. In particular if u is a coordinate function, it is trivial that u on a walk sample sequence cannot be convergent unless u on the sequence is finally constant. Then almost every walk sample sequence must hit the boundary (where it sticks), as was to be proved.

If ξ is not a boundary point, there is an integer-valued random variable ν (the first hitting time of the boundary) such that x_ν is a boundary point, but x_j is not for $j<\nu$. The random variable ν is a stopping time relative to the sequence of minimal σ-algebras. If u is harmonic on S, if $\nu_1=0$ and $\nu_2=\nu$, the stopping time theorem (slightly extended) asserts that $u(\xi)$, $u(x_\nu)$ is a martingale with two random variables. But then

$$u(\xi) = E\{u(\xi)\} = E\{u(x_\nu)\}. \tag{12.1}$$

Denote by $\mu(\xi, \eta)$ the probability that $x_\nu=\eta$, that is, the probability that a walk starting at ξ hits the boundary at η. The distribution $\mu(\xi, \cdot)$ is called **harmonic measure** relative to ξ. In terms of harmonic measure, (12.1) becomes

$$u(\xi) = \sum u(\eta)\mu(\xi, \eta), \tag{12.2}$$

where the sum is over all boundary points η. Thus a harmonic function on S is determined by its values on the boundary, in fact, is the weighted average using

harmonic measure of its values on the boundary. The rules for manipulating conditional expectations yield the fact that if u is an arbitrary function defined on the boundary of S, and if u is defined at interior points by (12.2), then u is harmonic on S. We have now shown the existence and uniqueness of a harmonic function with a specified boundary function. We omit the corresponding treatment of more general harmonicity defined using weighted averages, not necessarily at nearest neighbors. When $d=1$, the result obtained is particularly intuitive. In this case if S is the set of points $-a, \cdots, b$, where a and b are strictly positive integers, the walk is the random walk in which a step is either 1 or -1, with probability 1/2 each, independent of previous steps, until $-a$ or b is reached. In gambling language: a gambler is playing a fair game in which at each play he can win or lose a dollar with probability 1/2 each and the plays are independent; he starts with a dollars, his opponent with b dollars, and the game ends when he or his opponent has lost all his money; x_n is the gambler's total winnings (positive or negative) after the nth play. Since the play starts at time 0, we define $x_0=0$. If we take the harmonic function u to be the identity function, $u(\xi)=\xi$ on S, the sequence $\{x_n, n\geqq 0\}$ is seen to be a martingale, so (12.1) becomes

$$E\{u(x_\nu)\} = 0, \tag{12.1'}$$

that is, the expected final gain is 0 as it should be. Equation (12.1′) can also be obtained as a special case of (10.1). However obtained, this equation yields the standard result that the probability the gambler wins, that is, the probability that x_ν is b, is $a/(a+b)$.

13. Continuation of Section 7. If $d=2$, the properties of the random walk of Section 7 are intimately related to the properties of harmonic and analytic functions. We shall see this first in a simple application where S is the whole plane. It can be shown in this case that almost every sample walk starting from any point ξ is dense in the plane. This fact corresponds to Liouville's theorem that a bounded complex function which is analytic on the plane is a constant function, and we shall now prove Liouville's theorem probabilistically. Since the real and imaginary parts of an analytic function are harmonic, we shall obtain a stronger result if we prove that a function harmonic and bounded on the plane is constant. In fact we shall do even better and prove that a harmonic function on the plane which is bounded from above or below is a constant. By linearity we can suppose that the function u is positive, and we consider the martingale $\{u(x_n), n\geqq 0\}$. Trivially,

$$E\{u(x_0)\} = E\{u(x_n)\} = E\{\,|\,u(x_n)\,|\,\}.$$

According to the Martingale Convergence Theorem, this martingale is almost surely convergent. If we choose any sample walk which is dense and on which u has a limit, say c, the function u must be identically c by continuity, as was to be proved.

We proceed to the analog of the work in Section 12, assuming from now on that S is bounded and d arbitrary. The almost sure convergence of the bounded martingale $\{u(x_n), n \geqq 0\}$, when u is a coordinate function, implies that $\lim_{n\to\infty} x_n = x_\infty$ exists almost surely. This is impossible unless x_∞ has its values on the boundary of S. If ξ is the initial point of the walk and if A is a Borel boundary set, let $\mu(\xi, A)$ be the probability that x_∞ is in A. Then $u(\xi, \cdot)$ is a measure of boundary sets, *harmonic measure relative to* ξ. If u is a bounded harmonic function on S, or even if u is merely bounded above or below, the Martingale Convergence Theorem implies that $\lim_{n\to\infty} u(x_n)$ exists almost surely. That is, u has a limit at the boundary of S along almost every one of the sample walks. We shall continue this aspect of the discussion in Section 16. Suppose now that u is actually defined and continuous on the closure of S. Then trivially $\lim_{n\to\infty} u(x_n) = u(x_\infty)$ almost surely and it is straightforward to show that the ordered set of random variables

$$u(x_0), u(x_1), \cdots, u(x_\infty)$$

is a martingale. Since the expectations of the first and last random variables are equal,

$$u(\xi) = E\{u(x_\infty)\} = \int u(\eta)\mu(\xi, d\eta), \tag{13.1}$$

where the integration is over the boundary of S, that is, $u(\xi)$ is the average of its values on the boundary, weighted by harmonic measure. Conversely, if u is a function defined and continuous on the boundary of S, and if u is defined on S by (13.1), it can be shown, using the ideas in Section 16, that u is harmonic on S and has the assigned boundary function value as a limit at each boundary point near which the boundary is well-behaved (more precisely at each regular boundary point in the potential theoretic sense). Thus martingale theory solves the first boundary value problem for harmonic functions. This kind of analysis is applicable not only to classical harmonic functions, but also to the solutions of general elliptic and parabolic partial differential equations.

14. Example: application to derivation. The close relation between martingale theory and derivation theory is illustrated by the following example. Let Ω be the unit interval $[0, 1]$ and let the probability of a subset A be its Lebesgue measure, denoted by $|A|$. For each $n \geqq 1$, let $A_{n1}, \cdots, A_{nk_n}$ be a partition of $[0, 1]$ into disjoint intervals, and suppose that the $(n+1)$th partition is a refinement of the nth, that is, we suppose that each A_{nj} is a union of sets in the $(n+1)$th partition. If $\mathfrak{F}_n$ is the class of unions of sets in the nth partition, then $\mathfrak{F}_n \subset \mathfrak{F}_{n+1}$. If f is an integrable function on $[0, 1]$, define the random variable x_n by

$$x_n = \int_{A_{nj}} f(\omega)d\omega / |A_{nj}| \qquad \text{on } A_{nj}, \quad j \geqq 1,$$

to get a martingale relative to $\{\mathfrak{F}_n, n \geqq 1\}$ for which

$$E\{\,|\,x_n\,|\} \leqq \int_0^1 |\,f(\omega)\,|\,d\omega. \tag{14.1}$$

If f is continuous and if $\lim_{n\to\infty}\max_j|A_{nj}| = 0$, it is trivial that $\lim_{n\to\infty}x_n(\omega) = f(\omega)$ for all ω. Without this restriction on f and on the partitions the Martingale Convergence Theorem implies that $\lim_{n\to\infty}x_n$ exists almost everywhere on $[0, 1]$. More generally, $[0, 1]$ can be replaced by any measure space $\{\Omega, \mathfrak{F}, P\}$ with $P\{\Omega\} = 1$. In this new context the partition is to be a partition of Ω into countably many disjoint measurable sets; the $(n+1)$th partition is to be a refinement of the nth; $\int_A f d\omega$ is replaced by $\mu(A)$, where μ is any finite signed measure on $\mathfrak{F}$; x_n is defined as $\mu(A_{nj})/P\{A_{nj}\}$ on A_{nj} if the denominator does not vanish, defined arbitrarily on A_{nj} if $P\{A_{nj}\} = 0$. With this definition the sequence of random variables is a martingale if

(a) $P\{A_{nj}\} > 0$ for all n, j,

or if

(b) $\mu(A_{nj}) = 0$, whenever $P\{A_{nj}\} = 0$.

Without either (a) or (b) the sequence of random variables is a supermartingale if

(c) $\mu \geqq 0$.

Under (a) or (b) or (c), $E\{|x_n|\}$ is at most the absolute variation of μ, and we conclude that the martingale or supermartingale converges almost surely. Since a finite signed measure is the difference between two finite measures, we have derived the fact that a finite signed measure has a derivative relative to any finite measure with respect to a net of partitions. (The extension to allow $P\{\Omega\}$ to be other than 1 is trivial.)

15. Example: functions on an infinite dimensional cube. The definition of a martingale implies that each martingale random variable is a partial averaging of the next one. The following example exhibits this fact very neatly. Let Ω be the coordinate space of two way infinite sequences $\cdots, \xi_{-1}, \xi_0, \xi_1, \cdots, 0 \leqq \xi_j \leqq 1$. Let P be the product measure on Ω for which each coordinate measure is Lebesgue measure, that is, P is infinite dimensional volume. Let f be a measurable integrable function on Ω. Then we define, in the obvious notation, the random variable x_n by

$$x_n = \int\int \cdots f(\cdots, \xi_0, \xi_1, \cdots) \prod_{n+1}^{\infty} d\xi_j. \tag{15.1}$$

Here we have integrated out $\xi_{n+1}, \xi_{n+2}, \cdots$. If $\mathfrak{F}_n$ is the smallest σ-algebra of subsets of Ω containing every set determined by a condition of the form $\xi_j < c$ for $j \leqq n$, it can be shown that $x_n = E\{f \mid \mathfrak{F}_n\}$ almost surely, as should be intuitively clear: we are calculating the expected value of f knowing all the coordinates up to and including the nth. Then (see Sections 3 and 9) the two way sequence $\{x_n, -\infty < n < \infty\}$ is a martingale which converges in both directions: $\lim_{n\to\infty}x_n = x_\infty$ and $\lim_{n\to-\infty}x_n = x_{-\infty}$ exist almost surely. These limits can be shown to be (almost surely) f and $E\{f\}$ respectively.

16. Continuous parameter case. Let $\{x_t, t \in I\}$ be a supermartingale, where I is an interval on the line. Then Principle P_2 is illustrated by the fact that each random variable $\omega \rightsquigarrow x_t(\omega)$ can be changed on an ω set of probability 0 in such a way that almost every sample function (that is for almost every ω the function $t \rightsquigarrow x_t(\omega)$) becomes continuous except for nonoscillatory discontinuities. Such a change does not affect the joint distributions of finite sets of the random variables or the fact that the family of random variables is a supermartingale. The theory of continuous parameter supermartingales is very rich and plays an essential role in probabilistic potential theory. We give only one application here, a continuation of Sections 7 and 13. Let $\{z_t, 0 \leqq t < \infty\}$ be Brownian motion in d-dimensional space, with initial point ξ: z_0 is identically ξ; for each $t \geqq 0$ the random variable z_t has its values in d-space; every sample function $t \rightsquigarrow z_t(\omega)$ is continuous. We omit the exact specification of the distributions. It turns out that, in the notation of Section 7, if $T_0 = 0$, if T_1 is the first time t when z_t is a point of $B(\xi)$, if T_2 is the first time $t > T_1$ when z_t is a point of $B(z_{T_1})$ and so on, and if x_n is defined as z_{T_n}, then the sequence $\{x_n, n \geqq 0\}$ is precisely the walk discussed in Sections 7 and 13. Thus a walk sample sequence is a sequence of points on a Brownian path. If S is the whole space and if u is a real harmonic function (or, if $d = 2$, a complex analytic function) on S and if $|u|$ is not too large (we omit the precise restriction), then $\{u(z_t), t \geqq 0\}$ is a martingale. Suppose from now on that S is bounded. (The exactly appropriate condition is more generally that the complement of S has zero capacity.) Let $T = \sup_n T_n$ be the first time t at which z_t is on the boundary of S. Then it can be shown that T is almost surely finite, and of course z_T can be identified with x_∞ (defined in Section 13). Then the distribution of z_T is harmonic measure relative to the initial point of the Brownian motion. A refinement of the analysis in Section 13 shows that if u is harmonic and positive on S, and if $u(z_t)$ is defined as 0 when $t > T$, the process $\{u(z_t), t \geqq 0\}$ is a supermartingale. (This result is valid even if u is only superharmonic and positive, excluding the parameter value 0 if $u(\xi) = \infty$, and the discussion here can be generalized correspondingly.) From this result it is then concluded, using the existence of left limits of supermartingale sample functions, that u has a limit at the boundary of S along almost every Brownian path: $\lim_{t \uparrow T} u(z_t)$ exists almost surely. The limit theorem we have obtained is a probabilistic generalization of the classical FATOU THEOREM that if S is a disk, then any positive harmonic function u on S has a limit along almost every radius: if the disk radius is a and if polar coordinates are used, $\lim_{r \to a} u(re^{i\theta})$ exists Lebesgue for almost every θ. (The result was extended later to balls of arbitrary dimensionality.) Note that Fatou's Theorem is much more special because in his theorem S is a ball. On the other hand Fatou's paths are undeniably more pleasant, or at least more traditional, and are individually identifiable. The probability theorem, however, when applied to a ball, can be used to deduce Fatou's theorem, and in fact when boundary limit theorems are extended to cover superharmonic functions and more general classes of functions it becomes clear that the probabilistic versions are intrinsic, not accidents of the geometry of the domains.

References

1. J. L. Doob, Stochastic Processes, Wiley, New York, 1953.
2. Paul A. Meyer, Probability and Potentials, Blaisdell, Waltham, Mass., 1966.

MATHEMATICAL FOUNDATIONS FOR MATHEMATICS

LEON HENKIN,[1] University of California, Berkeley

1. Introduction. Most mathematical papers deal with mathematics "in the small"—a few definitions, a few theorems, a few proofs. If the author has a modicum of boldness and compassion he may also include some account of the intuitive ideas from which these formal parts of his work were fashioned. This paper, however, will have a different character.

In wondering what subject to choose for this Chauvenet Symposium, I let my mind's eye wander over those areas of the foundations of mathematics in which I have worked or dabbled—completeness proofs, applications of logic to algebra, decision problems, infinitary logic, algebraic logic . . . somehow none of them seemed appropriate. I began to wonder why. Presently it seemed to me that the answer was bound up with what might be called the "sociological structure" of our contemporary American mathematical community.

[1] The point of view toward foundations developed here was first formulated by me in the IBM Lectures which I gave at Swarthmore in December, 1967. This viewpoint has been developed over an extended period, during which much of my work was supported by the National Science Foundation (most recently, Grant No. GP-6232X).

This paper constitutes the final paper prepared for the Chauvenet Symposium held at the U. S. Naval Academy in October 1969 (this Monthly, May 1970).

Professor Leon Henkin received his PhD at Princeton University in 1947 under the direction of Alonzo Church. His thesis included a proof of Gödel's completeness theorem for the predicate calculus which has since become the standard proof in almost every presentation of mathematical logic. In addition Professor Henkin developed the theory of cylindrification algebras which is an algebraic formulation of the theory of quantifiers. His principal work has been in the area of foundations and mathematical logic in which he has published many papers and is a recognized authority. He was awarded the Chauvenet Prize in 1964 for his paper "Are Logic and Mathematics Identical?" published in Science, 1962 (vol 138).

Professor Henkin was Fine Instructor and Jewett Fellow at Princeton from 1947 to 1949, having spent four previous years as a mathematician in industry. In addition he was a Fullbright Scholar in Amsterdam in 1954–55, a Visiting Professor at Dartmouth in 1960–61, and a Guggenheim Fellow and member of the Institute For Advanced Study in 1961–62, and a visiting Fellow at All Souls' College, Oxford in 1968–69. He taught at the University of Southern California and has been a member of the faculty of the University of California at Berkeley since 1953 where he has served as chairman twice. He has been Editor for the *Journal of Symbolic Logic* and served a three year term as President for the Association of Symbolic Logic. He has also been a member of the Council of the American Mathematical Society as well as active in CUPM. Besides his papers in foundations he is the author of "Retracing Elementary Mathematics," Macmillan, 1962, which indicates his keen interest in the teaching of mathematics. *J. C. Abbott.*

For the fact is that mathematicians who work in areas of foundations are considered by most other mathematicians to be somewhat "different." Those working in any of the central parts of algebra, geometry, or analysis are generally familiar with the most basic notions in each other's domains, and have at some point worked on elementary problems dealing with these notions. But foundations, along with other "new" mathematical areas such as statistics or computer science, and certain "old" areas such as celestial mechanics, is considered a domain of whose details most mathematicians may safely remain ignorant, as long as they know in a vague way what the subject is about. These "fringe" areas of mathematics each relate mathematics to some nonmathematical discipline; but while, in the case of statistics, computer science, or celestial mechanics, this outside discipline is accepted as a proper domain for the application of mathematics; in the case of foundations the outside discipline is philosophy—and this renders the area just a little bit suspect.

I believe that this state of affairs makes it difficult for a "foundationist" to write an ordinary "in the small" paper for a general mathematical audience, unless it deals with the most elementary concepts. I have decided, therefore, to try to write about the foundations of mathematics "in the large."

Actually the phrase "foundations of mathematics" has a meaning which seems to shift from one context to another. Sometimes it denotes a part of mathematics—a sort of beginning portion on which all the other parts are built. At other times this phrase refers to a commentary on mathematics from the outside—a sort of explanation of the significance of the work of mathematicians, couched often in metaphysical terms.

If we take the pragmatic viewpoint that the foundations of mathematics is that which is done by those who profess to work in this area, two things become apparent. First, the *extent* of the domain, the number and distance-from-one-another of the subjects pursued under the heading of foundations of mathematics, has increased enormously during the course of recent decades. And second, the center of mass of this profusion of material has been moving continuously, and at an accelerated rate, from philosophy into mathematics.

Such a transition is by no means unique to the foundations of mathematics. All the empirical sciences, in their time, have been located squarely within "natural philosophy"; only gradually did they detach themselves and become separate disciplines. The independent fashioning of such sciences as physics, chemistry, and biology has not meant that philosophers have lost all interest in these disciplines. Quite the contrary, philosophers remain interested spectators, and to some extent even participants, in their ongoing development. And so it is with the foundations of mathematics. Even though the mathematician is increasingly taking up the subject and transmuting it thereby into a corpus having the form of mathematical definitions, theorems, and proofs, there remains nevertheless a residue of inquiry beyond the borders of mathematics which is a fertile—and by no means neglected—area for philosophical inquiry.

The entrance of mathematicians as a significant and recognized component

of those working in the field of foundations is generally dated around the turn of the century. The ground had been prepared by the development of new parts of mathematics, beginning with non-Euclidean geometry in the first half of the 1800's, Boole's work on logic about 1850, Cantor's creation of a theory of sets from 1870 on, and including the work of such men as Dedekind, Schroeder, Frege, Weierstrass, Peano, and Zermelo. But it remained for other mathematicians, writing somewhat later, to delineate and urge *viewpoints* concerning the foundations of mathematics, based upon these and further technical developments, which exerted intense and lasting influences on the further development of the subject. Best known among these were Kronecker, Poincaré, Brouwer, Hilbert, and Russell.

These men whose work in various areas of mathematics had in most cases already won them wide renown, engaged in a public debate concerning the nature and significance of mathematics, through a series of publications, and through letters which subsequently became public. Such fundamental controversy, almost unknown in the older branches of mathematics, naturally became the object of broad attention, and out of it there emerged three "schools" which have come to be known under the titles Intuitionism, Formalism, and Logicism.

Intuitionism, as Brouwer developed it from tendencies put forward by Kronecker and Poincaré, became a radical renunciation of the classical use of mathematical language and logic, and a demand for a severe limitation on the permissible kinds of mathematical constructions. Formalism, as Hilbert's program came to be known, focused attention sharply on the symbolic patterns of mathematical language, and attempted to explain their use in terms of rules having a syntactical character, whose ultimate justification was to be achieved by consistency proofs of so fundamental a character as to put them beyond dispute. Logicism, as Russell based its exposition on Frege's ideas, concentrated on the unification of mathematics, through systematic reduction of its parts to the most elementary and the most general—logic and the notion of sets.

In the six or seven decades since these positions were so ably set forth, mathematicians have taken up foundational work in steadily increasing volume. Much of the early material stemmed from, and fitted neatly into, the basic tripartite framework outlined above. But as in any lengthy mathematical development, continuous transformations, as well as abrupt changes resulting from completely unexpected developments, reshaped the discipline. New areas of foundational work appeared and new connections were found between foundations and other parts of mathematics.

Despite this burgeoning of mathematical activity, *commentators* who seek to classify and analyze the work have continued for the most part to try to fit the details into the logicism-formalism-intuitionism scheme. To us it appears that these efforts are increasingly unavailing, and lead to a more and more distorted view of what is being done. Perhaps the most basic reason is that the concepts embodied in these viewpoints are essentially philosophical in character, and as

such are not well adapted for describing material in accelerated transition from philosophy into mathematics.

Thus the time is at hand, it seems to us, to put forward a new framework, or classification scheme, within which cohesion, form, and meaning can be given to the totality of work in the foundations of mathematics. The concepts underlying such a classification scheme should be essentially mathematical in character, rather than philosophical, to take proper account of the increasingly mathematical nature of the work to be classified. The particular suggestion we shall advance here is that the continuing analysis of foundational work be based on the following classification scheme: set-theoretical aspects, algebraic aspects, and constructive aspects.

It is, of course, possible to consider that such a scheme is little more than a renaming of the same basic categories we have described earlier. Certainly the notions of set theory form a basic part of logicism, algebraic aspects of foundations can be closely related to the formalist's approach, and constructive elements of mathematics often arise from the same spirit of inquiry which motivated Brouwer to hoist the pennant of intuitionism. But though it is worthwhile to notice these connections, it seems to me that our proposed classification scheme differs from what has become the traditional one, in some fundamental ways.

One of the most basic differences between these two schemes has to do with the question of compatibility. Logicism, formalism, and intuitionism are essentially *competitive* ways of viewing the foundations of mathematics. Indeed, the original exchanges of publications and letters by the authors of these programs bordered on the acrimonious, because each considered that his was the *right* way to look at the foundations of mathematics, and that the others were (in consequence) *wrong*. By contrast, we find today a considerable cooperation among those pursuing set-theoretical, algebraic, or constructive elements in the foundations of mathematics. Each of these strands of the subject supplements the others; they are interwoven to provide a richly illuminated depiction of a common domain. Some of the most prized theorems and valuable insights are those which illuminate the interconnections among these strands of foundational work.

In the remainder of this paper we wish to illustrate the working of our proposed conceptual framework by indicating how some of the major complexes of problems which have occupied "foundationists" can be viewed within it.

2. Algebraic aspects of foundations. I still recall the definition of algebra given by my high school teacher when I took my first course in the subject: "It is like arithmetic," she said, "but we work with letters instead of numbers." Those who have pursued the subject beyond high school will be unable to give so simple an account of this part of mathematics, while agreeing that the use of letters as variables in equations remains an important characteristic of the subject.

One of the remarkable trends of recent mathematics has been the development of algebraic parts of various disciplines—we have algebraic topology, for example, algebraic geometry, functional analysis, and most recently algebraic logic. This incursion of algebra into other parts of mathematics represents, I believe, an important part of the effort to preserve a unity in mathematics, so that the development of more specialized disciplines may be counterbalanced by mathematical elements which facilitate communication among *all* mathematicians.

Although the term "algebraic logic" (coined by Halmos) is less than ten years old, the ideas of the subject are easily traced back along a continuous path to Boole's work in 1850. Indeed Boole's starting point was the observation that certain classical laws of logic could be expressed by means of algebraic equations; for example, the logical equivalence of a proposition with the denial of its denial took the form $--p=p$, while the law of the excluded middle was expressed by the identity $p+-p=1$. The fact that one could pass from such laws to others by the familiar algebraic methods of substitution for the variables of an equational identity, and the "replacement of equals by equals," served further to show an underlying identity between logic and the algebra of numbers.

Boole further enriched his theory by providing a second interpretation for his algebraic equations, in which variables represented sets instead of propositions. The equation $--p=p$ then came to mean that the complement of the complement of a set was the set itself, while $p+-p=1$ expressed the fact that the union of a set and its complement was the universal set. It remained for Stone, however, some eighty years later, to carry the interpretation of Boole's equations to the ultimate possibility by considering the class of *all* structures which satisfy them, calling these Boolean algebras, and studying their relations to one another in terms of such concepts as homomorphisms, subalgebras, and direct products which had evolved in the study of rings and groups.

Between Boole and Stone, however, there was other activity. Boole's calculus of classes led to an exhaustive study of a calculus of binary relations by Schroeder; Peirce took important steps toward bringing quantifiers within the scope of algebraically-expressed logic; and Tarski initiated a study of deductively closed systems of sentences which anticipated several of Stone's findings and methods.

The use of Boole's equations to define the class of Boolean algebras, and the central role played in their theory by Stone's representation theorem, had a profound effect on the further development of algebraic logic. One of the first of these was the establishment of a theory of relation algebras by Tarski: he selected some of the equational identities discovered by Schroeder in his study of binary relations, and used them as axioms to define a new class of algebras. He then studied the possibility of isomorphically representing arbitrary algebras of this class by means of binary relations over some set. This inquiry had an unusual history, as we shall relate below.

In studying the set of all binary relations over some set U one distinguishes

not only the empty relation, $\emptyset$, and the universal relation $U\times U$, but also the identity relation I_U. And one studies not only the Boolean operations $\cup$, $\cap$, and $\sim$ (complementation with respect to $U\times U$, under which the set of binary relations on U is closed), but also the operation of forming the converse $R^{\cup}$ for any given $R\subseteq U\times U$, and the binary operation of forming the relative product $R|S$ for given R and S, where for any x, $y\in U$ we have $x(R|S)y$ iff xRz and zSy for some $z\in U$. The Boolean identities form a trivial part of the theorems of this theory. Many of the identities involving I_u, $^{\cup}$, and $|$, are also simple: For example the associative law for $|$, the fact that I_u is a two-sided identity element for the operation $|$, the distributive law for $^{\cup}$ over $\cup$, the fact that $|$ is additive (i.e., that $(R_1\cup R_2)|S=(R_1|S)\cup(R_2|S)$ and $R|(S_1\cup S_2)=(R|S_1)\cup(R|S_2)$), and the law $(R|S)^{\cup}=S^{\cup}|R^{\cup}$. On the other hand, there are many equations involving these notions for which it is very difficult to determine whether they hold identically for relations over an arbitrary set U. Indeed, Tarski has shown that there does not exist a decision method which can determine automatically, in a finite number of steps, whether any given equation is such an identity. He does this by showing how an arbitrary statement of elementary number theory (a theory known to be undecidable) can be translated into an equivalent statement having the form of an equational identity for relation algebras. By contrast, it is known that a Boolean equation holds identically in any Boolean algebra iff it holds identically for the algebra of subsets of a one-element set, which of course provides a simple decision method for selecting the Boolean identities from among all Boolean equations.

After Tarski had selected certain identities from the theory of relations and used them as axioms to define the class of relation algebras, other identities were found, by Lyndon, which could not be derived from these axioms. It was natural to think of adding these new identities as further axioms, in an effort to obtain a system of axioms sufficiently strong to imply all identities of the theory of binary relations or, better still, strong enough to allow the inference that any structure satisfying the axioms must be isomorphic to an algebra of relations over some set. However, Lyndon stated that the latter goal was unattainable by presenting a relation algebra not isomorphic to any algebra of binary relations over a set U, which seemed to satisfy all the same equational identities as a certain algebra of the latter kind. Later, however, Tarski demonstrated a contrary result: He showed that any model satisfying the set of all equational identities which hold in every algebra of relations, must be isomorphic to such an algebra. (This he did by employing Garrett Birkhoff's criterion for equational classes of structures, showing that the class of isomorphic images of algebras of relations is closed under formation of subalgebras, homomorphic images, and direct products.) The mistake in Lyndon's proof was later found by Dana Scott. Still later, Monk showed that no finite system of identities can characterize the isomorphic images of algebras of relations.

A different approach to the theory of binary relations was taken by Everett and Ulam, who formulated the notion of a projective algebra. In place of the

operations of converse and relative product, they consider the projections P_0 and P_1 of an arbitrary relation $R \subseteq U \times U$ onto the "lines" $\{e\} \times U$ and $U \times \{e\}$, where e is a distinguished element of U; the Boolean operations on relations are retained in these structures. This approach algebraizes a much more limited part of the theory of relations than Tarski's, and the authors are able to establish a representation theorem for the class of projective algebras. It has also been shown that there is a decision procedure to determine automatically, in a finite number of steps, whether a given equation in this theory holds identically for arbitrary binary relations (or—equivalently, in view of the representation theorem—for arbitrary elements in any projective algebra).

Although the notion of projective algebra led to a very restricted theory, it suggested to Tarski, and his students Chin and Thompson, certain modifications which have led to a much richer development. The most significant change was to abandon a restriction to binary relations over a set U, and to consider relations of rank α over U for arbitrary α. Next, instead of considering the projections of a relation on some special "line" (or linear subspace of higher dimension for relations of higher rank), the fundamental operations are chosen to be the cylindrifications C_κ for each $\kappa < \alpha$: For any relation R of rank α over U, the relation $C_\kappa R$ is the set of all "points" $(x_0, \cdots, x_{\alpha-1})$ of ${}^\alpha U$ such that

$$(x_0, \cdots, x_{\kappa-1}, y, x_{\kappa+1}, \cdots, x_{\alpha-1}) \in R$$

for some $y \in U$. Finally, in case $\alpha = 2$ the identity relation over U is taken as a distinguished element just as in the case of relation algebras and, more generally, for any α we distinguish the "diagonal relations" $D_{\kappa\lambda}$ of rank α over U, where $(x_0, \cdots, x_{\alpha-1}) \in D_{\kappa\lambda}$ iff $x_\kappa = x_\lambda$ (for each $\kappa, \lambda < \alpha$). A set of relations of rank α over U is called a *cylindric field of dimension* α if it contains the relations $\varnothing$, ${}^\alpha U$, and each $D_{\kappa\lambda}$ (for $\kappa, \lambda < \alpha$), and if it is closed under the Boolean operations $\sim$ (complementation with respect to ${}^\alpha U$), $\cup$, $\cap$, and all of the cylindrifications C_κ (for $\kappa < \alpha$). The notion of a *cylindric algebra of dimension* α is obtained by abstracting from the notion of a cylindric field of relations; such structures have the form

$$\langle A, +, \cdot, -, 0, 1, d_{\kappa\lambda}, c_\kappa \rangle_{\kappa,\lambda<\alpha}$$

where $\langle A, +, \cdot, -, 0, 1\rangle$ is an arbitrary Boolean algebra, $d_{\kappa\lambda} \in A$ and c_κ is a one-place operation on A for each $\kappa, \lambda < \alpha$, and in which certain equations are satisfied identically. These equations, collected into seven axiom schemata, are of course chosen from among those equations known to hold identically in every cylindric field. Among these we may cite:

AXIOM C_3. $c_\kappa(x \cdot c_\kappa y) = c_\kappa x \cdot c_\kappa y$ *for all* $x, y \in A$.

AXIOM C_7. $c_\kappa(d_{\kappa\lambda} \cdot x) \cdot c_\kappa(d_{\kappa\lambda} \cdot -x) = 0$ *for all* $x \in A$, *where* $\kappa, \lambda < \alpha$ *and* $\kappa \neq \lambda$.

The feature that gives to cylindric algebras a special importance for logical studies is the fact that the cylindrifications c_κ stand in exactly the same relation to the existential quantifiers $(\exists v_\kappa)$ of predicate logic, as the Boolean operations

$+$, $\cdot$, $-$ bear to the connectives $\vee$ (disjunction), $\wedge$ (conjunction), and $\neg$, (negation) of sentential logic. Also the diagonal elements $d_{\kappa\lambda}$ correspond algebraically to the equations $v_\kappa = v_\lambda$ of predicate logic with equality. Thus the study of these algebras provides a way to deal with questions of predicate logic by algebraic methods. However, before we can be satisfied that the algebraic structure is adequate for a full logical analysis there are certain questions which must be considered.

In predicate logic, reflecting the needs of many parts of mathematics, we generally consider systems in which several relations of different ranks may be considered. How do we deal with these in the context of a cylindric algebra of dimension α, where we have one fixed value of α? Very simply, we choose α to be ω (the first transfinite ordinal), and for each integer κ we represent a relation of rank κ as a relation of rank ω which depends on only the first κ coordinates. For instance, if we take U to be the set of all real numbers and we wish to represent the binary relation $<$ on U as a relation of rank ω, we would use the relation R such that, for every infinite sequence $(x_0, x_1, \cdots)$ of real numbers we have $(x_0, x_1, \cdots) \in R$ iff $x_0 < x_1$.

In this example, we notice that if $\kappa \geqq 2$ and $(x_0, x_1, \cdots, x_\kappa, \cdots) \in R$ for some sequence $(x_0, x_1, \cdots, x_\kappa, \cdots)$ of real numbers, then if we change the coordinate x_κ in any way the resulting sequence is also in R; in other words we have $C_\kappa R = R$. More generally, whenever we represent a relation of rank λ as a relation R of rank ω, we shall have $C_\kappa R = R$ for every $\kappa \geqq \lambda$. Thus, in an arbitrary cylindric algebra A of dimension ω, if we consider the set A_λ of all those $x \in A$ for which $c_\kappa x = x$, whenever $\kappa \geqq \lambda$, we shall be dealing with an algebraic version of a cylindric field of relations of rank λ; it is not surprising, therefore, that the set A_λ together with the operations $c_0, \cdots, c_{\lambda-1}$ and the diagonal elements with indices $<\lambda$ forms a cylindric algebra of dimension λ.

Of course there are relations of rank ω which do not depend on only a finite number of coordinates, and hence which do not correspond to any relation of finite rank. In ordinary predicate logic, and most parts of mathematics, we deal only with relations of finite rank. Hence, if we wish to consider cylindric algebras which correspond precisely to systems of predicate logic, we impose the condition that for every $x \in A$ there is some $\lambda < \omega$ such that $c_\kappa x = x$ for all $\kappa \geqq \lambda$. Such a cylindric algebra is said to be *locally finite*.

Finally, let us consider the operation of forming the converse of a binary relation, one of the fundamental notions of relation algebras. In predicate logic the corresponding operation is a substitution, enabling us to pass from a formula Gxy to the formula Gyx. How can we obtain a corresponding operation in the theory of cylindric algebras? To see this, consider first the operation S_2^0 on relations, such that $S_2^0 R = C_0(D_{02} \cap R)$ for each R. If R is a relation of rank ω which represents a binary relation, so that $C_\kappa R = R$ for all $\kappa \geqq 2$, then we shall have

$$(x_0, x_1, x_2, x_3, \cdots) \in S_2^0 R \qquad \text{iff } (x_2, x_1, x_0, x_3, \cdots) \in R.$$

For the relation thus obtained we have $C_\kappa(S_2^0R)=S_2^0R$ for $\kappa=0, 3, 4, 5, \cdots$. Next, applying the operation S_3^1 to S_2^0R (where, in general, $S_3^1T=C_1(D_{13}\cap T)$ for any relation T), we find that

$$(x_0, x_1, x_2, x_3, \cdots) \in (S_3^1S_2^0R) \quad \text{iff } (x_0, x_3, x_2, x_1, \cdots) \in (S_2^0R)$$
$$\text{iff } (x_2, x_3, x_0, x_1, \cdots) \in R,$$

and $C_\kappa(S_3^1S_2^0R)=S_3^1S_2^0R$ for $\kappa=0, 1, 4, 5, 6, \cdots$. Continuing in this way by applying successively the operations S_0^3 and S_1^2 we finally obtain the relation $T=S_1^2S_0^3S_3^1S_2^0R$ such that $C_\kappa T=T$ for each $\kappa\geqq 2$ and

$$(x_0, x_1, x_2, x_3, \cdots) \in T \quad \text{iff } (x_1, x_0, x_2, x_3, \cdots) \in R.$$

Clearly, then, T is a relation of rank ω representing the converse of the binary relation represented by R.

Similarly, in an arbitrary cylindric algebra of dimension α we can define an operation S_λ^κ for each κ, $\lambda<\alpha$ so that $S_\lambda^\kappa x=C_\kappa(d_{\kappa\lambda}\cdot x)$ if $\kappa\neq\lambda$ and $S_\kappa^\kappa x=x$. If $\alpha=\omega$ and the algebra is locally finite, we can find a suitable finite sequence of these operations S_λ^κ to effect any given substitution on relations of given rank. For example, suppose S' is the substitution on relations such that

$$(x_0, x_1, x_2) \in S'R \quad \text{iff } (x_2, x_0, x_0) \in R$$

for any relation R of rank 3. Then S' corresponds to the operation $S_0^3 \circ S_0^1 \circ S_2^0 \circ S_3^2$, in the sense that for any relation R of rank ω such that $C_\kappa R=R$ for all $\kappa\geqq 3$, we shall have also $C_\kappa(S_0^3S_0^1S_2^0S_3^2R)=(S_0^3S_0^1S_2^0S_3^2R)$ for all $\kappa\geqq 3$ and

$$(x_0, x_1, x_2, x_3, \cdots) \in (S_0^3S_1^0S_2^0S_3^2R) \quad \text{iff } (x_2, x_0, x_0, x_3, \cdots) \in R$$

for each sequence $(x_0, x_1, x_2, x_3, \cdots)$.

We have thus seen how a locally finite cylindric algebra of dimension ω is equipped to deal with relations of any finite rank λ, and how its fundamental operations of cylindrification can be combined with its diagonal elements to permit the definition of an arbitrary operation of substitution on relations of given rank λ. Thus all of the elements to be found in a system of first-order predicate logic have their algebraic counterparts in cylindric algebras of this type.

For the class of locally finite cylindric algebras of dimension ω, there is the following representation theorem: Given any such algebra A, and any of its elements $x\neq 0$, there exists a homomorphism h of A into a suitable cylindric field relation, such that $hx\neq 0$. (By considerations of a general algebraic character, one can show that this is equivalent to the statement that every algebra A of the class considered is isomorphic to a subdirect product of cylindric fields.) This representation theorem may be considered as an algebraic analogue of the completeness theorem for predicate logic. Indeed, the two theorems are closely related, each of them being directly deducible from the other.

If we consider the class of all cylindric algebras of dimension ω, abandoning the condition of local finiteness, the representation theorem no longer holds. The class of arbitrary cylindric algebras of dimension ω includes cylindric fields of relations of rank ω which do not represent relations of any finite rank λ, i.e., which do not satisfy any condition of the form $C_\kappa R = R$ for all $\kappa \geqq \lambda$. Such algebras correspond to infinitary systems of predicate logic in which atomic formulas may consist of a predicate symbol followed by an infinite sequence of individual symbols. The failure of the representation theorem for the class of all ω-dimensional cylindric algebras means that there is such an algebra A which is not isomorphic to any subdirect product of cylindric fields; this is shown by producing an equation which holds identically in every cylindric field of dimension ω, but fails in A. In consequence we can find, in a system of infinitary predicate logic such as described above, a set of formulas which is consistent under the ordinary rules of deduction but cannot be satisfied in any model.

Of course we can consider the class of all cylindric algebras of *any* given dimension α, finite or infinite. Also, for any α we can consider cylindric fields whose elements are relations of rank α over some set U; and so we can ask whether all α-dimensional cylindric algebras are representable. It turns out that in each case the answer is negative, except for $\alpha = 0$ and $\alpha = 1$. In the case $\alpha = 2$, we can produce two equational identities which characterize the representable 2-dimensional cylindric algebras; for any $\alpha > 3$, no finite set of equational identities characterizes the set of representable α-dimensional cylindric algebras, although an α-dimensional cylindric algebra satisfying all equational identities which hold in every cylindric field of dimension α *is* always representable.

The class of locally finite algebras has a generalization: For any infinite α, we may consider those α-dimensional cylindric algebras A such that, for every $x \in A$, we have $c_\kappa x = x$ for at least one $\kappa < \alpha$. Such an algebra is said to be *dimension-complemented*, and every such algebra turns out to be representable. Here are some other classes of cylindric algebras known to be representable for any dimension α: (i) Every algebra in which the unit element can be expressed as a finite sum of diagonal elements $d_{\kappa\lambda}$ $(\kappa \neq \lambda)$; (ii) every algebra which can be generated by monadic elements, i.e., by elements x such that $c_\kappa x = x$ for all $\kappa > 0$; (iii) every algebra A of dimension α which can be isomorphically imbedded in an algebra B of dimension $\alpha + \omega$, in such a way that for every $x \in A$ we have $c_\kappa x = x$ for $\kappa = \alpha, \alpha + 1, \alpha + 2, \cdots$.

There are a great many open questions in the theory of cylindric algebras. Some of these are purely algebraic in character; for example, the question whether every such algebra which is finitely generated and simple (admits only trivial homomorphisms) can be generated by a single element: this is known to have a negative answer for cylindric algebras of finite dimensions, but is open for the infinite dimensional case. Other open problems relate arbitrary cylindric algebras with cylindric fields of relations. Of these, two problems occupy a central position: (i) To find, for arbitrary α, an intrinsic, algebraic characteriza-

tion of the class of all α-dimensional cylindric fields of relations; (ii) to find set-theoretic operations with intuitive geometric content, which can be performed on α-dimensional cylindric fields to yield a class of relational structures to provide an isomorphic representation of every cylindric algebra of dimension α. These questions have foundational significance because they are algebraic formulations of problems about a broad class of logical systems and their interpretations.

To complete this sketch of algebraic logic we mention that the concept of cylindric algebras is only one among several classes of algebraic structures that have been introduced in connection with logical systems. Certain classes are obtained by modifying the Boolean laws in order to deal with "nonclassical" logics. Among types of algebras employed for studying classical logic, polyadic algebras are those which, along with cylindric algebras, have been studied most intensively. Polyadic algebras are Boolean algebras whose fundamental structure is enriched by cylindrifications and by arbitrary substitutions (corresponding to the permutation and identification of individual symbols in predicate logic). The class of locally finite, ω-dimensional, polyadic algebras provides an algebraic theory of predicate logic without equality; algebras of this class which contain a set of diagonal elements are equivalent to locally-finite ω-dimensional cylindric algebras. For finite α, or for infinite α without the condition of local finiteness, the theories of these two classes have essential differences. In particular, every infinite-dimensional polyadic algebra can be represented by polyadic fields of relations.

We have been writing in some detail about the development of algebraic logic, for that is the most overt way in which algebraic ideas have entered into work on the foundations of mathematics. However, a total inventory of the algebraic aspects of foundational research would include many, many other areas. We shall mention three of these.

1. The whole development of deductive logic may be viewed as a process of algebraization. In order to understand this, let us sketch the common elements of the formation of almost any algebraic theory—theories of groups, say, or of rings, or of vector spaces. (i) We start with some particular domain of elements and operations on it. (ii) We develop a symbolic language to talk about this structure and we find sentences of the language, true of the structure, which seem to be related to one another by simple formal rules of transformation—for example, we may concentrate on sentences having the form of equational identities, related to one another by rules of substitution and replacement. (iii) We abstract from the particular domain in which we were first interested by selecting certain of the true sentences encountered in (ii) and using them as axioms, and we then study the class of all structures which satisfy those axioms. This is the heart of the process of algebraization. (iv) We look for representation theorems which relate arbitrary models of our axioms to the particular structure which gave rise to the theory. . . . This four-part process of development can be discerned in any well-developed algebraic theory.

Now how can we view deductive logic from this point of view?

(i) The most basic object of study in logic is the relation of implication. This is a relation which connects a given sentence (the conclusion) with a given set of sentences (the assumptions). Of course underlying this relation is a grammatical structure $\mathcal{L}$ within which the sentences are constructed—it consists of a list of symbols (classified into different kinds such as variables, connectives, parentheses, etc.) and rules for combining them into categories such as terms, formulas, and sentences. Once the grammar $\mathcal{L}$ is set forth in precise terms, it is necessary to describe how its components are *interpreted* in order to obtain a discourse language for some given structure (or other domain of discourse); under such an interpretation each sentence of $\mathcal{L}$ takes on a definite truth value, truth or falsity. Finally, a sentence is called *logically valid* if it is true under *every* interpretation; a sentence ϕ is a *logical consequence* of a set Γ of sentences, and Γ is said to *imply* ϕ, if ϕ is true in *all those* interpretations which make every sentence of Γ true.

(ii) To talk about the set of valid sentences and the relation of implication, we introduce variables like "ϕ" and "ψ" to range over sentences of $\mathcal{L}$, and variables like "Γ" and "Δ" to range over sets of sentences, and we introduce a special symbol "$\models$". We write $\models\phi$ to indicate that ϕ is valid; we write $\Gamma\models\phi$ to indicate that Γ implies ϕ. (This double use of the same symbol is reasonable since $\models\phi$ iff $\varnothing\models\phi$.) We also have symbols for the operations of building complex sentences from simpler ones, or from formulas, such as the familiar notations $\phi\rightarrow\psi$ for an "if . . . then" sentence or "$\forall v\phi$" denoting a "for all . . . " sentence. Finally, we find simple transformation rules allowing us to pass from certain expressions of the form "$\Gamma\models\phi$" which are true about implication, to others. For instance, the rule of detachment tells us that whenever we have $\Gamma\models\phi$ and $\Gamma\models\phi\rightarrow\psi$, we shall also have $\Gamma\models\psi$; another rule allows us to pass from $\Gamma\models\forall x(\phi\rightarrow\psi)$ to $\Gamma\models(\forall x\phi)\rightarrow(\forall x\psi)$.

(iii) Instead of the particular relation $\models$ we now abstract and consider an *arbitrary* relation, $\vdash$, satisfying certain of the conditions which we found to be satisfied for the particular relation $\models$. For example, if we have noticed that $\models\phi\rightarrow(\psi\rightarrow\phi)$ for all sentences ϕ, ψ of $\mathcal{L}$, we may impose the axiom $\vdash\phi\rightarrow(\psi\rightarrow\phi)$ for all ϕ, ψ of $\mathcal{L}$ on our undefined relation $\vdash$. Similarly, we may adopt the rule of detachment as an axiom on $\vdash$, requiring that whenever $\Gamma\vdash\phi$ and $\Gamma\vdash\phi\rightarrow\psi$, we also have $\Gamma\vdash\psi$ . . . . The passage from the theory of $\models$ to that of $\vdash$ is thus quite analogous to the passage from the study of the integers under addition to the consideration of the class of additive groups. But in this case, the theory of $\vdash$ at which we arrive is precisely what we call a formal deductive system. Those formulas θ for which we have postulated $\vdash\theta$ are called the *formal axioms* of the deductive system; those postulates about $\vdash$ having the form "whenever $\Gamma_1\vdash\theta_1$ and $\Gamma_2\vdash\theta_2$ then $\Gamma_3\vdash\theta_3$" define what we call the *formal rules of inference* of the deductive system. By combining formal axioms with formal rules of inference we obtain derived statements of the form $\vdash\theta$, e.g., we may derive $\vdash\phi\rightarrow\phi$ for all sentences ϕ of $\mathcal{L}$; such sentences θ are called *formal theorems*, and

the sequence of steps leading from axioms and rules to such a formal theorem is called a *formal proof*. Similarly, we may derive statements of the form "If $\Gamma_1 \vdash \theta_1$ then $\Gamma_2 \vdash \theta_2$," e.g., with formal axioms and rules of the usual kind we obtain the result that whenever $\Gamma \vdash \phi$ then also $\Gamma \vdash \psi \rightarrow \phi$. Such statements are called *derived rules of inference*. The derivation of such rules and of formal theorems constitutes the elementary part of deductive logic; it is entirely analogous to the derivation of identities holding in every group, from the group axioms.

(iv) Finally, we consider the totality of relations $\vdash$ (called *formal implications*) satisfying the postulates describing the formal axioms and rules of inference, and we seek to relate them to the relation $\models$ of logical inference from which we started. In the usual systems of elementary logic, for example, we are able to show that $\models$ is the intersection of all formal implications $\vdash$; we call this the *completeness* property of the deductive system, for it allows us to infer that whenever $\Gamma \models \phi$ then also $\Gamma \vdash \phi$ for every formal implication $\vdash$, and hence we can establish that $\Gamma \vdash \phi$ by using our formal axioms and formal rules of inference. The establishment of such a completeness result for a formal deductive system can thus be seen as a kind of representation theorem when the formulation of the deductive system is viewed as a process of algebraizing a theory of logical implication.

2. Another area of foundational work in which the algebraic aspects play an important role is in the theory of models. In this work we seek relations between the syntactical form of a given set of sentences, and the totality of structures which make all of these sentences true, i.e., which are models of the given set of sentences. Typically, these classes of models are characterized in terms of closure under algebraic operations for combining given structures to obtain new ones.

One of the earliest and most important examples of this type of work was the characterization of equational classes by Garrett Birkhoff. Suppose, for example, we are interested in structures of the form $(A, \circ, f)$, where A is a set closed under a 2-place operation $\circ$ and a 1-place operation f. (The totality of such structures is an example of a *similarity class*. This one we shall denote by S.) We consider a rudimentary language, $\mathfrak{L}$, equipped for discourse about any structure $(A, \circ, f) \in S$. $\mathfrak{L}$ is to contain variables $x, y, z, \ldots$ ranging over A, symbols $+$ denoting the operation $\circ$ and $-$ denoting the operation f, parentheses, and an equality sign. Using these symbols, we can form all sorts of equations, such as $x+y=y+x$, $-(x+y)=-y$, $--x=x+-y$, etc. Given any such equation, and any structure $(A, \circ, f) \in S$, we can inquire whether the equation holds identically in the structure or whether it fails when some elements in A are assigned as values of the variables. Finally, given any set Γ of equations we can consider the totality S_Γ of all those structures of S in which every equation of Γ holds identically. A class of structures determined in this way by a set of equations is called an *equational class*. The problem for the theory of models is

to characterize the equational classes from among all subclasses of S. The solution found by Birkhoff is simple and elegant: A subclass of S is equational if, and only if, it is closed under formation of subalgebras, homomorphic images, and direct products.

These three methods of obtaining new structures from given ones first became familiar to mathematicians from the study of groups and rings, were then found of use in classifying other varieties of structures, and were finally recognized as having a natural setting in the general theory of algebraic structures (now considered a part of the foundations of mathematics). Although various older methods of combining algebraic structures have proved of use in the theory of models, this theory has also led to the consideration of new methods devised especially for its own ends. One of the most useful of these is the notion of an *ultraproduct* of a given family of structures.

Suppose we again study the similarity class S of structures $(A, \circ, f)$ considered above, and fix attention on some particular family $\{(A_i, \circ_i, f_i)\}_{i\in I}$ of them. The *direct product* of this family, as is well known, is the following structure $(B, \circ, f)$:

(a) B is the set of all functions ϕ, with domain I, such that $\phi_i \in A_i$ for all $i\in I$;
(b) for any $\phi, \psi\in B$, the element $\phi \circ \psi$ of B is such that $(\phi\circ\psi)_i = \phi_i \circ_i \psi_i$ for all $i \in I$;
(c) for any $\phi\in B$, the element $f\phi$ of B is such that $(f\phi)_i = f_i\phi_i$ for all $i\in I$.

An ultraproduct of the family $\{(A_i, \circ_i, f_i)\}_{i\in I}$ is a homomorphic image of the direct product $(B, \circ, f)$ obtained in a special way, as follows.

We consider filters on I. These are sets F of subsets of I, such that F is closed under $\cap$, and that $X\in F$ and $X\subseteq Y\subseteq I$ implies $Y\in F$. If a filter I is proper, i.e., $\varnothing\notin I$, then it can always be extended to a filter I' such that for *every* $X\subseteq I$ either X or its complement is in F, but not both. A filter I' of this kind is called an *ultrafilter*. Choosing any ultrafilter F on I we can pass from the direct product $(B, \circ, f)$ of the family $\{(A_i, \circ_i, f_i)\}_{i\in I}$, to a new structure $(B/F, \circ/F, f/F)$ as follows. We consider the relation $\equiv_F$ on B such that for any $\phi, \psi\in B$ we have $\phi\equiv_F\psi$, iff $\{i:\phi_i=\psi_i\}\in F$. This is easily seen to be an equivalence relation, and we take B/F to be the set of equivalence classes ϕ/F for all $\phi\in B$. Next, we find that whenever $\phi\equiv_F\psi$ we also have $(f\phi)\equiv_F(f\psi)$; hence f induces an operation f/F on B/F such that $(f/F)(\phi/F)=(f\phi)/F$ for each $\phi/F\in B/F$. Similarly, whenever $\phi_1\equiv_F\psi_1$ and $\phi_2\equiv_F\psi_2$, then we also have $(\phi_1\circ\phi_2) \equiv_F(\psi_1\circ\psi_2)$; hence we obtain an operation $\circ/F$ on B/F such that

$$(\phi/F)(\circ/F)(\psi/F) = (\phi\circ\psi)/F$$

for any $\phi/F, \psi/F\in B/F$. A structure $(B/F, \circ/F, f/F)$ obtained in this way from the direct product $(B, \circ, f)$ and any ultrafilter F on I is called an *ultraproduct* of the family $\{(A_i, \circ_i, f_i)\}$.

Ultraproducts are of interest in the theory of models when we pass from the

equational language $\mathcal{L}$ for our similarity class S to the first-order language $\mathcal{L}'$ for S. In $\mathcal{L}'$ we form sentences by first combining equations of $\mathcal{L}$ by (repeated) use of the connectives $\neg$ (not), $\wedge$ (and), $\vee$ (or), $\rightarrow$ (if . . . then), $\leftrightarrow$ (iff), and then prefixing to the resulting formula Q a string of quantifiers, $\forall v$ (for all v) or $\exists v$ (for some v), one for each variable v occurring in Q. A class S_σ consisting of all structures $(A, \circ, f)$ of S for which some given sentence σ of $\mathcal{L}'$ is true, is called an *elementary class*; an *elementary class in the wider sense* is a subclass S_Γ of S consisting of all structures satisfying some set Γ of sentences of $\mathcal{L}'$.

Now the basic connections between ultraproducts and sentences of $\mathcal{L}'$ are as follows:

(i) For any sentence σ of $\mathcal{L}'$, any family $\{(A_i, \circ_i, f_i)\}_{i\in I}$ of structures of S, and any ultrafilter F of I, σ will be true of the ultraproduct $(B/F, \circ/F, f/F)$ of the family iff

$$\{i \in I: \sigma \text{ is true of } (A_i, \circ_i, f_i)\} \in F.$$

(ii) A subclass $T\subseteq S$ is elementary iff both T and $S\sim T$ are closed under formation of ultraproducts. T is elementary in the wider sense iff K is closed under ultraproducts and $S\sim T$ contains any ultraproduct of a family $\{(A_i, \circ_i, f_i)\}_{i\in I}$ such that $(A_i, \circ_i, f_i) = (A, \circ, f)$ for every $i\in I$, where $(A, \circ, f)$ is some element of $S\sim T$. (Such an ultraproduct is said to be an *ultrapower* of the structure $(A, \circ, f)$.)

(iii) Two structures $(A, \circ, f)$ and $(A', \circ', f')$ are called *elementarily equivalent* if there is no sentence σ of $\mathcal{L}'$ which is true of one and false of the other. A necessary and sufficient condition for two structures to be elementarily equivalent is that there exists an ultrapower of one which is isomorphic to an ultrapower of the other.

The result (i) can be proved very simply from the definition of ultraproducts. Results (ii) and (iii), due to Keisler, have been proved under the Generalized Continuum Hypothesis. It is an outstanding open problem of the theory of models whether this hypothesis can be eliminated from the proof.

3. The last area of foundational work we shall mention in which algebraic aspects are predominant, consists of applications of results or methods from foundational studies to obtain particular results in some part of algebra. One part of algebra in which such results have been obtained is the theory of real closed fields. In 1948 Tarski published a decision method to determine in a finite number of steps whether any given sentence in the first-order language for ordered fields is true of the field of real numbers. This method depends on a basic lemma to the effect that a first-order sentence is true of the field of reals iff it is true in every real closed field. Tarski himself indicated many applications of his result to particular problems in algebra; but it was Abraham Robinson who showed how Artin's solution of Hilbert's 17th problem could be obtained very simply from Tarski's lemma: Every polynomial p which is definite (takes only nonnegative values) can be expressed as a sum of squares of rational

functions whose coefficients lie in the field generated by the coefficients of p. By appealing to the completeness theorem of first-order logic this result was strengthened by showing that the number of squares needed to represent a given definite polynomial depended only on its degree and the number of its variables.

The most widely known application of foundational methods to the solution of an algebraic problem is the work of Ax and Kochen on the fields Q_p (p-adic completion of the rationals). Artin conjectured that every form of degree d over Q_p, in which the number of variables exceeds d^2, has a nontrivial zero in Q_p. This conjecture had been established for $d=2$ and 3. Ax and Kochen showed that the conjecture is true for arbitrary d with the possible exception of a finite set of primes p (depending on d). Subsequently, it was found that the original conjecture is not true in full generality. The work of Ax and Kochen involves the use of ultraproducts and the study of elementary classes in the similarity class of rings. Subsequently Ax turned some of these techniques to the study of finite fields and was able to settle an outstanding foundational problem: He showed the existence of a decision method to determine automatically, in a finite number of steps, whether any given first-order sentence holds for all finite fields.

3. Set-theoretic aspects of foundations. The explicit study of sets in relation to the foundations of mathematics was begun independently by two men, Cantor and Frege, at about the same time and place—Germany in the 1870's. Cantor's principal achievement was to develop set-theory as a foundation for the study of infinite sets, generalizing the notions of cardinal and ordinal numbers to apply to them. Frege, proceeding downward instead of upward, aimed to explain the theory of natural numbers in terms of set-theoretic notions.

The work of both these men foundered on the paradoxes. Burali-Forti showed in 1897 that there is a pair of incomparable ordinals, just about the time that Cantor published a proof of the contrary result. Frege's fundamental work was in press in 1902 when he received a letter from Russell indicating how contradiction could be developed in his system by considering the set of all sets which are not elements of themselves, and inquiring whether this set is an element of itself.

Both the Burali-Forti and the Russell paradoxes may be traced to an unbridled use of the "comprehension principle," which states that to any condition expressed by a sentential formula $\sigma(x)$ with one free variable, we may associate a set G consisting of all those objects x which satisfy the formula. By taking $\sigma(x)$ to be "x is an ordinal," we are led to Burali-Forti's paradox; by taking $\sigma(x)$ to be "x is a set which is not an element of itself" we are led to Russell's.

How are we to regard these paradoxes? They have something in common with the schoolboy riddle: "What happens when an irresistible force meets an immovable body?" Of course the answer is that there cannot be both an irresistible force and an immovable body: We can tell that much just from the

meaning of the words "irresistible" and "immovable." The question whether either such a force or such a body exists is one requiring empirical investigation or physical theory to answer. . . . In set theory we have seemingly irresistible forces, such as the operation of passing from a given set G to the set $G \cup \{G\}$, which seems to indicate that every set can be enlarged. On the other hand, the set of all sets appears like an immovable body, incapable of enlargement because of its comprehensive character. Seeing clearly that we cannot have both, which one do we reject?

The way in which such fundamental decisions were made was through the formulation of axiomatic theories of sets. The version most widely known was presented by Zermelo in 1908. In this version the irresistible force is firmly ensconced, since each set G can be enlarged to $G \cup \{H\}$, where H is a subset of G for which we can prove $H \notin G$. The immovable object is not present, the argument of Russell's paradox being used to show that no set can have all sets among its elements. Still, the theory is intended to be comprehensive, as it claims to be a theory of all sets whatever. Zermelo's system was clarified by Skolem and strengthened by Fraenkel. The resulting system, ZF, is widely employed.

In the same year, 1908, Russell published his theory of types. Couched in the language of logic (propositional functions, rather than sets, are the basic objects of study), and encumbered by a hierarchical theory of ramified types to which a mysterious axiom of reducibility was added, it was hardly recognized as a theory of sets. As subsequently clarified and simplified by Ramsey and Chwistek, however, the theory of types emerges as a theory of sets—much more modest in scope, however, than Zermelo's. In its simplest version it deals with a sequence of domains $D_0, D_1, \ldots$, where D_0 is the domain of "individuals" (nonsets) and D_{n+1} is the domain of all subsets of D_n for each natural number n. Of course there is no pretense that these domains contain all sets whatever, but there are enough sets to develop the theories of the classical number systems and geometric spaces, and indeed, just about all parts of mathematics except a full theory of transfinite cardinal and ordinal numbers. The theory of types, like ZF, seems to favor the irresistible force over the immovable body: given any set in one of the domains D_n, we can always find larger ones in D_{n+1}; there is no all-inclusive set.

An axiomatic theory in which an immovable body is present to resist any enlarging force was originally conceived by von Neumann, recast by Bernays, and put in final form by Gödel; we call it the vNBG system. In it the basic objects are called classes, and two sorts are distinguished: those which are elements of other classes and those which are not. (The former are called sets, the latter proper classes.) Thus there is a class V of all sets, but it cannot be enlarged to $V \cup \{V\}$ since V is not an element of any class and so $\{V\}$ cannot be formed.

Although both the systems ZF and vNBG are set forth as though they deal with all sets whatever, there is a way of looking at these axiom systems which makes each of them appear, like the theory of types, as furnishing a theory for

other in which the axioms can be enumerated in an automatic manner, are known from Gödel's incompleteness result to possess sentences which are true but unprovable from the axioms. Furthermore, these sentences are of a very simple kind—indeed by Matijasevitch's recent solution of Hilbert's tenth problem we now know that they can be put in the form of sentences asserting the non-existence of a solution to some given diophantine equation. However, if we pass from a given theory of sets to one with sets of higher rank, we can generally prove some of these diophantine statements which were unprovable in the original system.

While diophantine statements have a simple form, they tend to lose interest when they contain 50 or 60 variables. By contrast, there are some quite short sentences which are known to be undecidable in any of the systems discussed above (type theory, ZF, vNBG)—namely, the axiom of choice and the continuum hypothesis, proved consistent by Gödel in 1936 and independent by Cohen in 1963. Several mathematicians have shown that the addition of various axioms asserting the existence of larger ordinals will not result in a system in which these axioms become decidable. (Incidentally, Cohen's independence results have been re-derived by Scott and Solovay using a notion of Boolean-valued models of set theory in which algebraic and set-theoretic methods are beautifully combined for foundational studies.)

So far we have been discussing axiomatic theories of sets, but this is by no means the only part of foundational studies we intend to classify under "set-theoretical aspects." Let us mention two others.

1. Set theory and logic are closely interrelated. We have described above how deductive logic may be considered as an algebraization of the notion of logical consequence. But how is this relation originally defined? Since $\Gamma \models \sigma$ holds iff the sentence σ is true in every structure which satisfies each sentence of Γ, we are led to inquire under what conditions a given sentence is true for a given structure. This inquiry was shown by Tarski, in 1934, to have no precise answer for sentences of natural language, which is inherently paradoxical; but he gave a mathematical definition of truth for sentences of formalized mathematical languages by using basic notions of set theory. This definition, and the more general semantical notion of satisfaction of a sentential formula by a sequence of elements of the given structure, have played an important role in many subsequent investigations into undecidable theories, relative consistency proofs, and the comparative strengths of two given theories.

The basis of any language is a grammar, G, possessing various kinds of symbols from which sentences and other syntactical categories can be built up. The syntactical operations on formulas of G allow us to consider the grammar as a kind of algebraic structure. These structures are "free" in their similarity class. When we interpret the grammar G with respect to a mathematical structure S, we are essentially mapping the structure G homomorphically—not into S itself, but into a cylindric field of relations over S, generated by the funda-

mental operations and relations of S. Thus the set-theoretical notions involved in cylindric fields of relations, as well as those required to justify recursive definitions in free structures, are bound up with the definition of truth, and hence with the fundamental logical notion of consequence. Although, for first-order languages, this semantically-defined relation can be shown equivalent to a syntactically defined notion of derivability, it is known from Gödel's work on incompleteness that this duality cannot be extended to higher-order languages. Thus the semantical notions, rooted in set theory, provide our only access to the logic of higher-order languages.

2. In 1959, in Warsaw, an international symposium was held on infinitistic methods in logic. A few years before, logicians had begun to study the logic of formal languages differing radically from natural languages in that they contained infinitely long sentences. Sentences of such languages could be infinitistic in three ways:

(i) atomic sentences could consist of a relation symbol of transfinite rank followed by a transfinite sequence of individual symbols,

(ii) infinite sets of sentences could be combined into one by conjunction, $\sigma_1 \wedge \sigma_2 \wedge \sigma_3 \wedge \ldots$, or by disjunction, $\sigma_1 \vee \sigma_2 \vee \sigma_3 \vee \ldots$, and

(iii) quantification could be carried out over infinitely many variables arranged in arbitrary order type, e.g.,

$$\cdots \forall v_6 \exists v_4 \forall v_2 \exists v_1 \forall v_3 \exists v_5 \cdots \phi(v_1, v_2, \cdots).$$

At first these infinitistic languages were regarded as curiosities, but in 1960 they were used by Tarski and Hanf to solve an old and difficult problem of set theory —thereby opening a new wave of research based on axioms of infinity which carry us much farther into the transfinite than any earlier ones.

The problem involves the existence of measures on sets of various cardinality. We shall be concerned with measures m, defined on all subsets of a set A, taking only values 0 and 1, such that $m(A)=1$ and $m(X)=0$ for each one-element set $X \subseteq A$. If κ is a cardinal, m is called κ-additive if $m(\bigcup_{i\in I} X_i) = \sum_{i\in I} m(X_i)$ whenever the family $\{X_i\}_{i\in I}$ of subsets of A has cardinality $<\kappa$. Of course if A is denumerable, i.e., of power ω, it has a measure m which is κ-additive for every $\kappa \leqq \omega$. For every nondenumerable cardinal α less than θ_1, the first nondenumerable inaccessible cardinal, it was known that a set A of cardinality α cannot have a κ-additive measure except for the case $\kappa \leqq \omega$. (In particular, it has no countably additive measure.) It was suspected that θ_1 would resemble the smallest inaccessible cardinal, admitting κ-additive measures for every $\kappa \leqq \theta_1$. Tarski, however, using Hanf's result about certain infinitary languages, was able to show that sets of cardinality θ_1 admit κ-additive measures only for $\kappa \leqq \omega$, hence have no countably additive measures.

Hanf's result deals with the compactness question of logic. For first-order languages of the ordinary kind, with sentences of finite length, the compactness theorem states that if Γ is a set of sentences such that the sentences of every

finite subset of Γ are true in some structure, then there is a model in which *every* sentence of Γ is true. Hanf considered infinitary languages L_α for each cardinal α, in which one can form the conjunction of any set of sentences whose cardinality is $<\alpha$, and one can quantify simultaneously over any set of variables whose cardinality is $<\alpha$. For each α satisfying $\omega<\alpha\leqq\theta_1$, he produced a set Γ of sentences of L_α such that Γ has cardinality α, the sentences of each subset of Γ of cardinality $<\alpha$ are all true in some structure, but there is no structure in which every sentence of Γ is true.

Hanf extended this result far into the transfinite. For example, if $\theta_0, \theta_1, \ldots, \theta_\xi, \ldots$ is the sequence of all inaccessible cardinals, then θ_1 in the above theorem can be replaced by any θ_ξ such that $\xi<\theta_\xi$. For any α less than such a θ_ξ Tarski's result also holds: A set of cardinality α admits no countably additive, 2-valued measure defined on all of its subsets.

These methods of proof seem to offer no hope of extending the result to *every* cardinal $\alpha>\omega$. Accordingly, various mathematicians have begun to explore the consequence of assuming, as a very strong axiom of infinity, that there is a cardinal $\alpha>\omega$ which is measurable, i.e., admits an α-additive measure. For example, Scott has found that Gödel's axiom of constructibility is incompatible with the existence of a measurable cardinal.

4. Constructive aspects of foundations. Intuitionism is generally regarded as the most radical constructivist position in the foundations of mathematics. Its radicalism consists in the rejection of large parts of mathematics which are accepted by most mathematicians. For instance, the law of the excluded middle, allowing us to assert for any statement p that either p or not-p must hold, is rejected in contexts where infinitely many objects are under discussion, except for particular statements p, where we have a method to decide whether p holds, or whether not-p holds. Again, no argument is accepted for a statement of the form "there is an integer p such that $\sigma(p)$," short of a proof which furnishes a specific way of arriving at a particular integer p such that $\sigma(p)$ can be established.

While it is not hard to ascribe these rejections to a desire for a more constructive approach to mathematics, it seems that the radical changes sought by intuitionists have other motivations, too. One of these is bound up with the name "intuitionism," and relates to the fundamental question of the significance of mathematical discourse. For the "classical" mathematician such discourse serves to communicate "facts" about certain abstract objects such as numbers, sets, and points. For the intuitionists, however, the aim of mathematics is "mental mathematical construction," and discourse is directed toward communication about this kind of intuitive experience. Such a disparity as to the purpose of mathematical discourse is sharply indicated by the following contention of some mathematicians about the intuitionist position: that a certain mathematical statement, e.g., about the existence of a number with a certain arithmetical property, may be true at certain times but not at others, or true

for some mathematicians but not for others. I myself am unsure whether intuitionists agree with this contention—perhaps some do and others don't. . . . At any rate it is certain that the intuitionistic and the classical mathematician use language in very different ways, and it seems evident from this fact that clear communication between them is not to be expected, unless and until they can agree on what constitutes a satisfactory translation from the language of one to that of the other.

One of the complaints by classical mathematicians about intuitionists is that the latter have been unwilling to commit themselves to clear-cut criteria for judging the correctness of a mathematical argument by furnishing an explicit list of axioms and rules of inference—i.e., a formal deductive system. As a gesture toward improving communication, Heyting—the leading exponent of intuitionism after Brouwer—published a paper in 1930 setting forth formal rules of intuitionistic propositional logic. All these rules are included among the classical laws of logic, but many of the latter are not derivable in Heyting's system. For instance, $p \rightarrow \neg\neg p$ is in, $\neg\neg p \rightarrow p$ is out. Heyting emphasizes, however, that while the rules of his system can be established intuitionistically, no formal system can be proved to exhaust the totality of laws which may be established by further intuitionistic constructions. It should also be noted that whereas classically the laws of logic come first, in that they are used in the development of other parts of mathematics, the intuitionist starts with arithmetic—constructions with the whole numbers—and only arrives at laws of logic as a kind of generalization of universally-valid methods of arithmetical construction.

Heyting's system of intuitionistic propositional logic has been the subject of various foundational studies pursued by means of classical mathematical methods. First Gödel showed that even though the Heyting system was on the face of it weaker than classical propositional logic, one could nevertheless associate with each propositional formula σ another one, σ', such that σ is classically valid if and only if σ' is provable in the Heyting system.

Tarski showed how a set-theoretical interpretation could be provided for the Heyting system. Classically, we associate with each formula σ of propositional logic a function σ_c of several arguments ranging over the subsets of an arbitrary set U. (The function σ_c is defined recursively; for example, for any given arguments the value of $(\neg\sigma)_c$ is the complement of the value of σ_c with respect to U.) Then a formula σ will be classically valid iff σ_c is the constant function with value U. In Tarski's interpretation, with each formula σ one associates a function σ_i whose arguments and values range over the open sets of an arbitrary topological space U. (For example, the value of $(\neg\sigma)_i$ is the interior of the complement of the value of σ_i, for any given arguments.) The formulas provable in Heyting's system are just those with constant value U in the Euclidean plane.

The most important methodological analysis of intuitionistic mathematics by classical methods was undertaken by Kleene. His tool was the recursive function, originally developed as a mathematically precise notion equivalent

to that of automatically calculable function, for the purpose of dealing with decision problems. Using this class of functions, Kleene defined a notion of realizability for sentences of arithmetic, which he proposed as an intuitionistic counterpart to the classical notion of truth. Kleene was able to show that the sentences realizable by some number were just those provable in Heyting's system of intuitionistic arithmetic. Attempts to extend this approach to intuitionistic logic were not as successful.

Recursive functions have been used in various ways to provide constructive alternatives to portions of classical mathematics—alternatives quite different from those afforded by intuitionism. For example, there is *recursive analysis*. The basic objects are real numbers treated as infinite decimal expansions, but one admits only those numbers whose decimal parts admit some rule by which one can recursively, i.e., automatically, compute the digit in the nth place, for any given n. Again, in dealing with real-valued functions of a real variable, one admits only those functions which can be computed in a recursive way for given arguments. This type of analysis, vigorously pursued a few years ago, has lost steam since the discovery that there was more than one way to limit the usable functions by recursive requirements that seemed intuitively plausible, but these ways were inequivalent.

There is also recursive set theory. In ordinary set theory two sets have the same cardinality when there is a one-one mapping of one onto the other. In the recursive version one deals only with sets of integers, and limits oneself to one-one mappings which are (partial) recursive. The resulting "cardinal numbers" are called isols, and their arithmetic has been extensively studied. . . . Another part of recursive set theory deals with recursive degrees of unsolvability. Where A and B are sets of natural numbers, we say that A is recursive in B if the question whether any given number is in A can be reduced by a recursive method to a finite number of questions about certain numbers being in B. A is recursively equivalent to B if each is recursive in the other; the corresponding equivalence classes are the degrees of unsolvability; the partial order induced on them by the relation A is recursive in B then determines a semi-lattice whose structure has been extensively studied.

It is interesting to note that though the study and application of recursive functions is rooted in finitistic approaches to mathematics, recent work has seen an impulse to extend this notion from functions on natural numbers to those which have a transfinite ordinal domain. Carried far enough, this theory makes contact with the universe of constructible sets created by Gödel for proving the consistency of the continuum hypothesis—a domain seemingly far removed from the finite.

5. Concluding remarks. Although we began by saying we would attempt to write a paper about the foundations of mathematics in the large, it should be very clear that we have not attempted to be all inclusive. Many important areas have gone unmentioned—we may cite the theories of categories, of

hierarchies, of decision problems, and automata, to name but a few. We have, however, tried to touch upon sufficiently many broad areas to give the reader some sense of the extent of the subject, and more particularly to illustrate our thesis that these materials can be roughly classified according to algebraic, set-theoretical, and constructive aspects of the subject. These aspects are not competitive, but supplement each other to illuminate fundamental problems from different perspectives. There are many interrelationships among the subjects we have mentioned; only a few have been made explicit, partly because of the author's ignorance and in part because these relations are still only obscurely understood. We hope that readers with varying mathematical interests may be led to clarify some of these questions.

Covering as much ground as we have done, it would be easy to append an enormous bibliography of related works. The reader who would like to see that kind of a list is invited to inspect volume 26 of the Journal of Symbolic Logic, given over entirely to a bibliography of work in that field during the preceding 25 years. Instead, we have chosen a very small list of works where the non-specialist reader may obtain more detail about some of the topics we have mentioned. The bibliographies of these works will then lead further into the subject for those who still have interest.

A limited number of copies of all the Chauvenet Symposium Papers bound as a single volume may be obtained by writing to Professor J. C. Abbott, Department of Mathematics, U. S. Naval Academy, Annapolis, Md., 21402.

References

1. J. W. Addison, L. Henkin, A. Tarski, International symposium on the theory of models, University of California, 1963. (North-Holland, Amsterdam, 1965).

2. J. Ax and S. Kochen, Diophantine problems over local fields: I, Amer. J. Math., 87 (1965) 605–630; II, *Ibid.*, vol. 87 (1965) 631–648; III, Ann. Math., 83 (1966) 437–456.

3. Paul Benacerraf, and Hilary Putnam, Philosophy of mathematics, selected readings, Prentice-Hall, Englewood Cliffs, 1964.

4. Paul J. Cohen, Set theory and the continuum hypothesis, Benjamin, New York, 1966.

5. K. Gödel, The consistency of the axiom of choice and the generalized continuum hypothesis with the axioms of set theory, fourth printing, Princeton University Press, Princeton, 1958.

6. Paul Halmos, Algebraic logic, Chelsea, New York, 1962.

7. L. Henkin, D. Monk, and A. Tarski, Cylindric algebras, vol. I, North-Holland, Amsterdam. (Forthcoming.)

8. A. Heyting, Intuitionism, an introduction. (Studies in logic, North-Holland, Amsterdam, 1956.)

9. Carol R. Karp, Languages with expressions of infinite length. (Studies in logic, North-Holland, Amsterdam, 1964.)

10. H. J. Keisler, and A. Tarski, From accessible to inaccessible cardinals, results holding for all accessible cardinal numbers and the problem of their extension to inaccessible ones, Fundamenta Mathematicae, LIII (1964) 225–308.

11. S. C. Kleene, Introduction to metamathematics, Van Nostrand, Princeton, N. J., 1952.

12. A. Lévy, and R. Solovay, Measurable cardinals and the continuum hypothesis, Israel J. Math., 5 (1967) 234–248.

13. R. Montague, and R. L. Vaught, Natural models of set theories, Fundamenta Mathematicae, XLVII (1959) 219–242.

14. Andrzej Mostowski, Thirty years of foundational studies, lectures on the development of mathematical logic and the study of the foundations of mathematics in 1930–1964, Acta Philos. Fenn., XVII (1965) 1–180.

15. Abraham Robinson, Introduction to model theory and to the metamathematics of algebra. (Studies in logic, North-Holland, Amsterdam, 1963.)

16. A. Tarski, A decision method for elementary algebra and geometry, University of California Press, Berkeley. 2nd ed., 1951.

17. ———, Logic, semantics, metamathematics, Papers from 1923 to 1938. (Translated by J. H. Woodger.) Clarendon Press, Oxford, 1956.

18. van Heijenoort, From Frege to Gödel, a source book in mathematical logic 1879–1931, Harvard University Press, Cambridge, Mass., 1967.

19. Infinitistic methods, proceedings of the symposium on foundations of mathematics, Warsaw 1959. Pergamon Press, New York, 1961.

20. Le raisonnement en mathématiques et en sciences expérimentales. Colloques internationaux du Centre National de la Recherche Scientifique, vol. LXX (1958).

21. Summaries of the talks presented at the Summer Institute for Symbolic Logic, Cornell University, 1957. (Second edition, I.D.A., 1960.)

PATTERNS OF VISIBLE AND NONVISIBLE LATTICE POINTS

FRITZ HERZOG and B. M. STEWART, Michigan State University

1. Introduction. Let L_k, for $k \geqq 2$, denote the k-dimensional lattice, i.e., the set of points $(x_1, x_2, \cdots, x_k)$ with integral coordinates x_λ. A point in L_k will be called *visible* (namely, visible from the origin) if and only if its components x_λ have no common divisor greater than 1. Otherwise, the point will be called *nonvisible*. (By this definition, the origin itself is counted among the nonvisible points, a choice which we make purely for the sake of convenience.) In our geometrical representations we shall frequently use the following symbols for the points of L_k:

visible point = circle (○), nonvisible point = cross (×).

As a mnemonic aid we note that the first vowel both in visible and circle is an "*i*"; the first vowel both in nonvisible and cross is an "*o*".

For instance, Fig. 1 gives the distribution of circles and crosses among the points (x, y) of L_2 with $0 \leqq x \leqq 10$, $0 \leqq y \leqq 10$.

By a *pattern* P_k we mean the following: to each of the w^k lattice points $(x_1, x_2, \cdots, x_k)$ with $1 \leqq x_\lambda \leqq w$ is associated either a circle or a cross or neither

Fritz Herzog received his Columbia University Ph.D. under J. F. Ritt. His first position was at Cornell, and he has been at Michigan State since 1943, except for visits to Washington University and the University of Michigan. He is a recipient of a Michigan State Distinguished Faculty award. His principal research interests are complex function theory and power series.

B. M. Stewart received his Wisconsin Ph.D. under C. C. MacDuffee and joined the Michigan State faculty in 1940. He was an MAA Governor from 1956–59 and is the author of *Theory of Numbers* (Macmillan 1952, 1964) and *Adventures among the Toroids—A Study of Orientable Polyhedra* (1970). His research interests are number theory, matrix theory, graph theory, and Euclidean geometry. *Editor.*

10	×	○	×	○	×	×	×	○	×	○	×
9	×	○	○	×	○	○	×	○	○	×	○
8	×	○	×	○	×	○	×	○	×	○	×
7	×	○	○	○	○	○	○	×	○	○	○
6	×	○	×	×	×	○	×	○	×	×	×
5	×	○	○	○	○	×	○	○	○	○	×
4	×	○	×	○	×	○	×	○	×	○	×
3	×	○	○	×	○	○	×	○	○	×	○
2	×	○	×	○	×	○	×	○	×	○	×
1	○	○	○	○	○	○	○	○	○	○	○
0	×	○	×	×	×	×	×	×	×	×	×
	0	1	2	3	4	5	6	7	8	9	10

FIG. 1

one of these two symbols. A typical example of a pattern P_2 is given in Fig. 2, where $w=6$ and where a dot signifies that neither a circle nor a cross has been associated with that point.

6	×	·	○	·	·	·
5	·	○	·	·	·	·
4	○	×	○	×	·	·
3	·	○	·	○	·	·
2	×	○	○	○	·	·
1	·	○	·	·	·	·
	1	2	3	4	5	6

FIG. 2

The question which we answer in this paper is: "Under what conditions (necessary and sufficient) can a given pattern P_k be 'realized' in L_k?" By this we mean: "Does there exist in L_k a hypercube of w^k points which has a visible point wherever P_k has a circle and a non-visible point wherever P_k has a cross? In other words, does there exist a point $(u_1, u_2, \cdots, u_k)$ in L_k such that the points

$$(u_1 + x_1, u_2 + x_2, \cdots, u_k + x_k) \tag{1}$$

are visible or nonvisible, whenever $(x_1, x_2, \cdots, x_k)$ is a point of P_k marked with a circle or cross, respectively?" It is clear that the points in (1) with $(x_1, x_2, \cdots, x_k)$ in P_k can be interpreted to be the result of a translation of

P_k, a terminology which we shall frequently employ. Also, it is evident that the choice of the hypercube $1 \leqq x_\lambda \leqq w$ for the pattern P_k is quite arbitrary; in fact, two patterns should obviously be considered as identical if one can be obtained from the other by a translation. It is equally evident that any rigid motion that maps L_k onto itself (including a reflection) leaves a realizable pattern realizable and vice versa.

The question asked in the preceding paragraph is answered for $k=2$ by Theorem 1 in Section 2 and for $k>2$ by Theorem 2 in Section 4. According to these theorems, the conditions that P_k can be realized in L_k are entirely independent of the distribution of the crosses in P_k and depend solely on the location of the circles. A simple example of a pattern P_k that cannot be realized in L_k is the pattern

$$\begin{matrix} \circ & \circ \\ \circ & \circ \end{matrix}$$

in the case $k=2$ and its k-dimensional equivalent for $k>2$. Evidently one of these 2^k points has all its coordinates even and is therefore nonvisible, no matter where in L_k these points may be located. In Section 3 we give a number of corollaries which state whether or not certain patterns P_2 can be realized. Also a number of actual realizations of a few patterns P_2 are given as examples.

It follows from the preceding remarks (see also Corollary 1, Section 3) that L_k contains arbitrarily large hypercubes consisting entirely of nonvisible points. This result seems almost to contradict our intuition, in view of the fact that the relative frequency of the visible points in L_k is $1/\zeta(k)$, so that the nonvisible points constitute the minority set in L_k (39% for $k=2$, 17% for $k=3$, 8% for $k=4$, etc.). For the case $k=2$ a proof of the above-mentioned fact concerning the relative frequency of the visible points can be found, for instance, in [**2**, pp. 105–107]. For the case $k>2$, see [**1**, Theorem 3], where however no proof is given. We supply here a brief proof, modeled after the one given by Rademacher in [**2**, *loc. cit.*]. It turns out that the case $k>2$ is actually simpler than the case $k=2$.

For integral $k \geqq 3$ and real $t \geqq 1$, let $\Psi_k(t)$ denote the number of visible lattice points $(x_1, x_2, \cdots, x_k)$ with $1 \leqq x_\lambda \leqq t$. Exactly as in [**2**, *loc. cit.*; see in particular p. 107, line 6], we obtain

$$\Psi_k(t) = \sum_{d=1}^{[t]} \mu(d) \left[\frac{t}{d}\right]^k = \sum_{d=1}^{\infty} \mu(d) \left[\frac{t}{d}\right]^k. \tag{2}$$

If in the last member of (2) we omit the brackets, we obtain the quantity

$$\sum_{d=1}^{\infty} \mu(d) \frac{t^k}{d^k} = \frac{t^k}{\zeta(k)} \cdot \tag{3}$$

We note that $0 \leqq x^k - [x]^k < kx^{k-1}$ for $x>0$. Hence the error $R_k(t)$ in omitting the brackets in (2) satisfies the inequality

$$
\begin{aligned}
|R_k(t)| &\leqq \sum_{d=1}^{\infty}\left\{\left(\frac{t}{d}\right)^k - \left[\frac{t}{d}\right]^k\right\} < \sum_{d=1}^{\infty} k\left(\frac{t}{d}\right)^{k-1} \\
&= kt^{k-1}\zeta(k-1) = o(t^k), \quad k \geqq 3.
\end{aligned} \tag{4}
$$

It follows from (2), (3), and (4) that $\Psi_k(t)/t^k \to 1/\zeta(k)$ as $t \to +\infty$.

Let S_k be a subset of L_k. We say S_k is *connected* if and only if any two points P and Q in S_k can be connected in S_k, i.e., there exists a finite chain of points $P_0 = P, P_1, P_2, \cdots, P_{h-1}, P_h = Q$ such that each P_i is in S_k and each distance $d(P_{i-1}, P_i) = 1$. We ask whether the set V_k of all visible points is connected and answer in the negative by the use of our Corollary 2, say, when $k=2$, and its analogue for $k>2$. For a similar result in a hexagonal plane tessellation, see [3]. In a subsequent paper the authors intend to investigate various problems concerning the connected components of visible points and those of nonvisible points, particularly in L_2.

2. The two-dimensional case. In this section we state and prove a necessary and sufficient condition for a given pattern P_2 to be realizable in L_2. To avoid subscripts we call the coordinates x and y instead of x_1 and x_2.

We shall need the notion of a *complete square modulo* m, where m is any positive integer. By this we mean a set S of m^2 points of L_2, say (x_ν, y_ν), $\nu = 1, 2, \cdots, m^2$, which form the Cartesian product of a complete system of residues modulo m with itself. This means, given any point (x, y) in L_2 with $0 \leqq x < m$, $0 \leqq y < m$, there is exactly one point (x_ν, y_ν) in S such that $(x_\nu, y_\nu) \equiv (x, y) \bmod m$. (Here and in what follows, the last congruence means that the relations $x_\nu \equiv x \bmod m$ and $y_\nu \equiv y \bmod m$ both hold.)

THEOREM 1. *A given pattern P_2 can be realized in L_2 if and only if the set C of circles in P_2 fails to contain a complete square modulo p for every prime p.*

REMARK. Referring to Fig. 2 in Section 1, we note that in that pattern C certainly cannot contain a complete square modulo p when $p \geqq 5$, since C consists of only ten points. Neither does C contain a complete square modulo 2, since no circle (x, y) is congruent to $(1, 1)$ modulo 2. C does, however, contain a complete square modulo 3, for instance, $C - \{(2, 5)\}$. Thus, according to Theorem 1, the pattern of Fig. 2 cannot be realized in L_2.

Proof. To prove the necessity of the condition, assume that we are given a pattern P_2 whose set C of circles contains a complete square modulo p for some prime p. Now choose an arbitrary point (u, v) in L_2 for the purpose of translating P_2. The set of points $(u+x, v+y)$ with (x, y) in C will also contain a complete square modulo p. Hence for at least one point (x, y) in C we shall have $(u+x, v+y) \equiv (0, 0) \bmod p$. This makes the point $(u+x, v+y)$ nonvisible, which violates the requirements for realizing the pattern P_2. The latter therefore cannot be realized.

To prove the sufficiency, we assume that we are given a pattern P_2 whose set C of circles satisfies the condition of Theorem 1. We proceed in three steps.

STEP 1. Let p be a prime with $p \leqq w$. We recall here that the entire pattern

P_2 is imbedded in the square $1 \leqq x \leqq w$, $1 \leqq y \leqq w$. By the condition of the theorem there exists a point, say (x_p, y_p), such that $(x, y) \not\equiv (x_p, y_p) \bmod p$ for all (x, y) in C. Let

$$(u, v) \equiv (-x_p, -y_p) \bmod p. \tag{5}$$

Then for all (x, y) in C we have $(u+x, v+y) \equiv (x-x_p, y-y_p) \not\equiv (0, 0) \bmod p$, so that none of the points $(u+x, v+y)$ with (x, y) in C will have both of its coordinates divisible by p.

Since the various moduli p in (5) are relatively prime in pairs, we may use the Chinese Remainder Theorem to choose u and v in such a way that (5) holds simultaneously for every prime $p \leqq w$.

STEP 2. To each cross (i, j) in the given pattern P_2 we associate a prime $Q(i, j) > w$, different primes $Q(i, j)$ to different points (i, j). To the congruence conditions (5) we then adjoin the set of congruences

$$(u, v) \equiv (-i, -j) \bmod Q(i, j), \tag{6}$$

where the (i, j) run over all crosses of P_2. We shall then have $(u+i, v+j) \equiv (0, 0) \bmod Q(i, j)$, so that the point $(u+i, v+j)$ will be nonvisible, as required for a realization of P_2, whenever (i, j) is a cross.

We again use the Chinese Remainder Theorem to satisfy both systems (5) and (6). In preparation for Step 3, we remark that since $(u+i, v+j) \equiv (0, 0) \bmod Q(i, j)$ and since $Q(i, j) > w$, the congruence $(u+x, v+y) \equiv (0, 0) \bmod Q(i, j)$ cannot hold for any point of the square $1 \leqq x \leqq w$, $1 \leqq y \leqq w$, other than (i, j); in particular, it cannot hold for any (x, y) in C.

STEP 3. At this stage the crosses of P_2 are entirely taken care of by Step 2; but the circles of P_2 are taken care of by Step 1 only as far as primes $p \leqq w$ are concerned. It is still possible at this point of our procedure that a prime greater than w might divide both coordinates of a point $(u+x, v+y)$, where (x, y) is in C. If such a prime exists, it is, by the remark made at the conclusion of Step 2, different from any of the primes $Q(i, j)$ chosen in Step 2. In order to rectify this situation we proceed as follows.

From the Chinese Remainder Theorem we know that u is determined from (5) and (6) uniquely modulo $\prod p \prod Q(i, j)$. We can therefore demand that u be positive. Let such a value of u be chosen and kept fixed from now on. The positive numbers $u+1, u+2, \cdots, u+w$ then have a finite number of prime factors greater than w and different from the $Q(i, j)$ of Step 2; for these prime factors we shall use the notation q. We now subject v, in addition to (5) and (6), to the further set of congruences

$$v \equiv 0 \bmod q, \tag{7}$$

where q runs over all primes selected above. We are sure that $v+y \not\equiv 0 \bmod q$ for all (x, y) in C and for all those primes q, because $q > w$ and therefore all of the numbers $v+1, v+2, \cdots, v+w$ lie between two successive multiples of q, namely, v and $v+q$. Hence for all (x, y) in C, although q may divide $u+x$, it will not divide $v+y$, which is the situation we desire.

Since the q in (7) are different from the p in (5) and from the $Q(i, j)$ in (6), we can once more appeal to the Chinese Remainder Theorem and satisfy the systems (5), (6), and (7) simultaneously. We note that, by our choice of the q in Step 3, a prime other than the p, the $Q(i, j)$ and the q cannot divide any of the numbers $u+1, u+2, \cdots, u+w$. The proof of Theorem 1 is therefore complete.

Before proceeding to the corollaries and examples of the next section, we make two remarks concerning the procedure described in the sufficiency part of the proof of Theorem 1. In the first place, this procedure makes it obvious that, if a given pattern P_2 can be realized in L_2, then it will occur in L_2 infinitely often. Secondly, the procedure in Steps 1, 2, and 3 was designed purely with the view in mind to make the proof as simple as possible. In the case of a given pattern which can be realized, simplifications in the procedure are numerous and varied, but also for the most part obvious, so that we need not go into any details here.

3. Corollaries and Examples. Each of the following Corollaries treats a type of pattern P_2 and is an immediate consequence of Theorem 1. The details of some of the proofs are left to the reader.

Each of the Examples covers a realizable pattern that was treated in the preceding Corollary. The abscissae and ordinates under and to the left of the pattern are to serve for better orientation. In the solution of the Example the values of these coordinates are given as products of their prime factors, so that the reader can tell at a glance whether a point is visible or nonvisible. The procedure we followed in obtaining our solution for each example can be easily reconstructed from the solution itself. For instance, it is clear that in Example 2 we used the following primes for the $Q(i, j)$:

$$\begin{array}{ccc} 2 & 11 & 2 \\ 3 & 17 & 5 \\ 2 & 7 & 2 \end{array}$$

COROLLARY 1. *Every pattern P_2 consisting only of crosses can be realized.*

Proof. The condition of Theorem 1 is vacuously satisfied.

The significance of Corollary 1 was discussed in the Introduction.

Example 1. The pattern

$$\begin{array}{ccc} v+2 & \times & \times \\ v+1 & \times & \times \\ & u+1 & u+2 \end{array}$$

is found at $u=13$, $v=19$. This gives

$$u+1 = 2\cdot 7, \qquad v+1 = 2^2\cdot 5,$$
$$u+2 = 3\cdot 5, \qquad v+2 = 3\cdot 7.$$

Example 2. The pattern

$$\begin{array}{cccc} v+3 & \times & \times & \times \\ v+2 & \times & \times & \times \\ v+1 & \times & \times & \times \\ & u+1 & u+2 & u+3 \end{array}$$

is found at $u=1307$, $v=1273$. This gives

$$\begin{aligned} u+1 &= 2^2\cdot 3\cdot 109, & v+1 &= 2\cdot 7^2\cdot 13,\\ u+2 &= 7\cdot 11\cdot 17, & v+2 &= 3\cdot 5^2\cdot 17,\\ u+3 &= 2\cdot 5\cdot 131, & v+3 &= 2^2\cdot 11\cdot 29. \end{aligned}$$

COROLLARY 2. *Every pattern P_2 consisting of one, two, or three circles and any number of crosses can be realized. In particular, there are "extremely lonesome" visible points in L_2 that are separated from all other visible points in L_2 by an arbitrarily great distance.*

Example 3. The pattern

$$\begin{array}{cccc} v+3 & \times & \times & \times \\ v+2 & \times & \circ & \times \\ v+1 & \times & \times & \times \\ & u+1 & u+2 & u+3 \end{array}$$

is realized by $u=53$, $v=19$, which gives

$$\begin{aligned} u+1 &= 2\cdot 3^3, & v+1 &= 2^2\cdot 5,\\ u+2 &= 5\cdot 11, & v+2 &= 3\cdot 7,\\ u+3 &= 2^3\cdot 7, & v+3 &= 2\cdot 11. \end{aligned}$$

Example 4. For the pattern

$$\begin{array}{ccccc} v+3 & \times & \times & \times & \times \\ v+2 & \times & \circ & \circ & \times \\ v+1 & \times & \times & \times & \times \\ & u+1 & u+2 & u+3 & u+4 \end{array}$$

choose $u=29753$, $v=15859$. We then find

$$\begin{aligned} u+1 &= 2\cdot 3^3\cdot 19\cdot 29, & v+1 &= 2^2\cdot 5\cdot 13\cdot 61,\\ u+2 &= 5\cdot 11\cdot 541, & v+2 &= 3\cdot 17\cdot 311,\\ u+3 &= 2^2\cdot 43\cdot 173, & v+3 &= 2\cdot 7\cdot 11\cdot 103.\\ u+4 &= 3\cdot 7\cdot 13\cdot 109, \end{aligned}$$

COROLLARY 3. *Let P_2 consist of the rectangle with vertices* $(1, 1)$, $(M, 1)$, (M, N) *and* $(1, N)$, $M \geqq 2$, $N \geqq 2$, *with all its "boundary" points being circles and all its "interior" points being crosses. Then P_2 can be realized if and only if M and N are both odd.*

Example 5. For $M=5$, $N=5$, that is, for the pattern

$$\begin{array}{cccccc}
v+5 & \circ & \circ & \circ & \circ & \circ \\
v+4 & \circ & \times & \times & \times & \circ \\
v+3 & \circ & \times & \times & \times & \circ \\
v+2 & \circ & \times & \times & \times & \circ \\
v+1 & \circ & \circ & \circ & \circ & \circ \\
 & u+1 & u+2 & u+3 & u+4 & u+5
\end{array}$$

choose $u=102$, $v=6198$, so that

$$\begin{array}{ll}
u+1 = 103, & v+1 = 6199, \\
u+2 = 2^3\cdot 13, & v+2 = 2^3\cdot 5^2\cdot 31, \\
u+3 = 3\cdot 5\cdot 7, & v+3 = 3^2\cdot 13\cdot 53, \\
u+4 = 2\cdot 53, & v+4 = 2\cdot 7\cdot 443, \\
u+5 = 107, & v+5 = 6203.
\end{array}$$

COROLLARY 4. *Let P_2 consist of the rectangle with vertices at* $(\pm m, \pm n)$, $m \geqq 1$, $n \geqq 1$, *with circles at* $(0, 0)$ *and at all the boundary points of P_2 and crosses at all interior points of P_2 other than the origin. This pattern P_2 can be realized if and only if* $6 \mid mn$.

COROLLARY 5. *Let P_2 consist of the "square diamond" with vertices at* $(\pm m, 0)$ *and* $(0, \pm m)$, $m \geqq 1$, *with circles along the edges of the square and crosses at the points interior to the square. This pattern P_2 can be realized for all values of m.*

Proof. We note that the circles of P_2 lie on the four lines $\pm x \pm y = m$. Let p be any prime. If $m \not\equiv 0 \bmod p$, then no circle of P_2 can be congruent to $(0, 0)$ modulo p; and if $m \equiv 0 \bmod p$, then no circle of P_2 can be congruent to $(1, 0)$ modulo p. The condition of Theorem 1 is therefore satisfied.

Example 6. Let $m=2$. The pattern

$$\begin{array}{cccccc}
v+2 & & & \circ & & \\
v+1 & & \circ & \times & \circ & \\
v & \circ & \times & \times & \times & \circ \\
v-1 & & \circ & \times & \circ & \\
v-2 & & & \circ & & \\
 & u-2 & u-1 & u & u+1 & u+2
\end{array}$$

has the solution $u=6$, $v=105$. This gives

$$\begin{aligned} u-2 &= 2^2, & v-2 &= 103,\\ u-1 &= 5, & v-1 &= 2^3\cdot 13,\\ u &= 2\cdot 3, & v &= 3\cdot 5\cdot 7,\\ u+1 &= 7, & v+1 &= 2\cdot 53,\\ u+2 &= 2^3, & v+2 &= 107. \end{aligned}$$

4. The higher-dimensional case. In this section we assume $k\geqq 3$. We define a *complete k-dimensional hypercube modulo m* in a way exactly corresponding to the case $k=2$ in Section 2. A complete k-dimensional hypercube modulo m has, of course, m^k points. We can then state the following result which generalizes Theorem 1 to any number of dimensions:

THEOREM 2. *A given pattern P_k can be realized in L_k if and only if the set C of circles in P_k fails to contain a complete k-dimensional hypercube modulo p for every prime p.*

Proof. The proof of Theorem 2 follows exactly the lines of the proof of Theorem 1 in Section 2 as far as the necessity part and Steps 1 and 2 of the sufficiency part are concerned. The system of congruences corresponding to (5) and (6) are here

$$(u_1, u_2, \cdots, u_k) \equiv (-x_{1p}, -x_{2p}, \cdots, -x_{kp}) \bmod p, \tag{5'}$$

$$(u_1, u_2, \cdots, u_k) \equiv (0, 0, \cdots, 0) \bmod Q(i_1, i_2, \cdots, i_k). \tag{6'}$$

In Step 3, we choose for u_1 a fixed positive value satisfying (5′) and (6′). After that, however, it suffices to subject only u_2 (and not $u_3, u_4, \cdots, u_k$) to the further system of congruences

$$u_2 \equiv 0 \bmod q, \tag{7'}$$

where q has the same meaning as it did in (7). This is because if a prime is not a common divisor of the first two coordinates of the point $(u_1+x_1, u_2+x_2, \cdots, u_k+x_k)$, it certainly cannot be a common divisor of all k of its coordinates.

Example. Let P_3 be the cube with vertices (i_1, i_2, i_3) where $i_\lambda=1$ or 2. All of the eight points of P_3 are to be crosses. As a solution we obtain $u_1=463198,9$ $u_2=2250169$, $u_3=2383093$, so that

$$\begin{aligned} u_1+1 &= 2\cdot 5\cdot 11\cdot 17\cdot 2477, & u_1+2 &= 3\cdot 7\cdot 13\cdot 19^2\cdot 47,\\ u_2+1 &= 2\cdot 5\cdot 13\cdot 19\cdot 911, & u_2+2 &= 3^2\cdot 7\cdot 11\cdot 17\cdot 191,\\ u_3+1 &= 2\cdot 7\cdot 17^2\cdot 19\cdot 31, & u_3+2 &= 3\cdot 5\cdot 11^2\cdot 13\cdot 101. \end{aligned}$$

References

1. John Christopher, The asymptotic density of some k-dimensional sets, this MONTHLY, 63 (1956) 399–401.

2. Hans Rademacher, Lectures on Elementary Number Theory, Blaisdell, New York, 1964.

3. Advanced Problem 5263, this MONTHLY: Proposal by Art Winfree, 72 (1965) 192–193; Solution by J. P. Altgeld, 73 (1966) 209–211. (The diagram accompanying 5263 has misleading errors: for (33, 9), (33, 22), (35, 15) and (35, 21) should be white, not black; and (35, 24) should be black, not white.)

ALGORITHMS INVOLVING ARITHMETIC AND GEOMETRIC MEANS

B. C. CARLSON, Iowa State University, Ames

1. Introduction. The Babylonian method of extracting the square root of a positive number a is to make a first guess, say x_0, and refine it successively by computing

$$x_{n+1} = \frac{1}{2}\left(x_n + \frac{a}{x_n}\right), \qquad n = 0, 1, 2, \cdots. \tag{1.1}$$

It is easy to show that $x_n \to \sqrt{a}$ as $n \to \infty$ and that the rate of convergence is quadratic, i.e., the error $x_{n+1} - \sqrt{a}$ is ultimately proportional to the square of $x_n - \sqrt{a}$.

The one-dimensional iterative algorithm can be rewritten in a two-dimensional form:

$$x_{n+1} = \frac{x_n + y_n}{2}, \qquad y_{n+1} = \frac{2}{\dfrac{1}{x_n} + \dfrac{1}{y_n}}, \qquad n = 0, 1, 2, \cdots, \tag{1.2}$$

by defining $y_n = a/x_n$ for each value of n. It follows that y_n also approaches the limit $\sqrt{a} = \sqrt{x_0 y_0}$, but a no longer appears in the recurrence relations. Note that x_{n+1} and y_{n+1} are the arithmetic and harmonic means, respectively, of x_n and y_n.

Iterative algorithms for computing special functions by means of recurrence relations of the form

$$x_{n+1} = f(x_n, y_n), \qquad y_{n+1} = g(x_n, y_n), \qquad n = 0, 1, 2, \cdots, \tag{1.3}$$

have received some attention in recent years because they are easy to use in modern computing machines. As Tricomi observed [**15**, pp. 13, 37–41], the two-dimensional mapping determined by f and g may have a continuum of fixed points, and there is then no general theory for determining the limits (if they exist) of x_n and y_n. Some interesting results have been found recently by Lehmer [**17**]. Only those rare cases are easy to handle in which one can discover an ex-

B. C. Carlson did his Ph.D. at Oxford under G. S. Rushbrooke. He spent four years in the Physics Department of Princeton University before joining the Physics Department at Iowa State; he ended up in the Mathematics Department where he has been except for a leave of absence to spend at Caltech. He is Fellow of the American Physical Society and his major fields of research are low-energy nuclear physics and special functions. *Editor.*

plicit invariant of the recurrence relations, such as the invariant $x_n y_n$ of the relations (1.2).

Two such tractable cases are Gauss' algorithm of the arithmetic-geometric mean, in which the harmonic mean in (1.2) is replaced by the geometric mean, and Borchardt's algorithm for computing an inverse sine. In both cases the functions f and g are constructed from arithmetic and geometric means. The present paper treats in a unified way all algorithms, including Gauss' and Borchardt's, in which $f(x, y)$ and $g(x, y)$ are chosen from the quantities $\frac{1}{2}(x+y)$, $(xy)^{1/2}$, $[\frac{1}{2}(x+y)x]^{1/2}$, and $[\frac{1}{2}(x+y)y]^{1/2}$. In each case we shall find an invariant in the form of an elliptic integral (usually a degenerate one such as an inverse sine or a logarithm), and from the invariant and the initial values, x_0 and y_0, we can determine the common limit of x_n and y_n.

Among these algorithms, Gauss' is unique in its quadratic convergence, which accounts for its continuing use in practical computation of elliptic integrals. It is unique also in its historical importance as the key to Gauss' private discovery of general elliptic functions at the end of the eighteenth century. It is not, however, unique in having historical associations, for another algorithm is a new and surprising postscript to an important discovery made by Fagnano in the early part of the eighteenth century. The algorithm provides an iterative solution to the problem of rectifying an arc of Bernoulli's lemniscate, a curve in the shape of a figure eight with equation $r^2 = \cos 2\theta$ in plane polar coordinates. Fagnano showed in 1718 how to double a lemniscatic arc with ruler and compass [**11**], [**16**], and the new algorithm provides a succinct demonstration of Fagnano's duplication theorem. Fagnano's result led Euler in 1753 to the addition theorem for lemniscatic arcs and thence to the addition theorem for general elliptic integrals. Because of this chain of events, Jacobi [**18**] described as "*einen für die Geschichte der Mathematik ungemein wichtigen Tag*" the day of December 23, 1751, when Euler received Fagnano's collected works in support of his nomination to the Berlin Academy. (The phrase "*Der Geburtstag der elliptischen Funktionen*" has sometimes been ascribed to Jacobi, perhaps incorrectly.)

Although elliptic functions and integrals got their names from the problem of rectifying an arc of an ellipse, this problem appears in retrospect to have been a red herring for the development of analysis in the eighteenth century. The really fruitful problem was the rectification of the lemniscate. Almost half a century after Euler's discovery of the addition theorem, Gauss inverted the problem by regarding the radial coordinate r as a function of the arc length. If r is assumed to change sign whenever it passes through zero, the function is qualitatively similar to the sine function and is called the lemniscatic sine [**9**], [**11**]. It is plainly periodic with a period equal to the perimeter of the lemniscate, and Gauss discovered that it is the restriction to the real axis of an analytic function that is doubly periodic in the complex plane, i.e., an elliptic function. He discovered also that the periods are connected with the arithmetic-geometric mean of 1 and $\sqrt{2}$ and subsequently that arithmetic-geometric means are in general connected with the periods of elliptic functions [**10**].

Thus each of the two major discoveries which grew from the lemniscate problem is related to an iterative algorithm. Gauss' discovery of general elliptic functions was tied both logically and historically to the algorithm of the arithmetic-geometric mean, which can be used with suitable initial conditions to compute the perimeter of the lemniscate. Fagnano's duplication theorem is a logical but not a historical consequence of an algorithm which rectifies a general arc of a lemniscate and the perimeter in particular.

2. Statement of four algorithms. The algorithm of the arithmetic-geometric mean occurred first in 1784–85 in a memoir by Lagrange [8] on reduction and evaluation of elliptic integrals. Gauss wrote that he conceived it independently in 1791, at age fourteen, and in 1799 it led him, as mentioned earlier, from lemniscate functions to general elliptic functions. In the statement of the algorithm x_0 and y_0 are positive numbers, and the common limit of x_n and y_n as $n \to \infty$, called the **arithmetic-geometric mean** of x_0 and y_0, is denoted by $L_{12}(x_0, y_0)$ for reasons to appear later:

$$
\begin{aligned}
&x_{n+1} = \frac{x_n + y_n}{2}, \qquad y_{n+1} = (x_n y_n)^{1/2}, \qquad n = 0, 1, 2, \cdots, \\
&\frac{1}{L_{12}(x_0, y_0)} = \frac{2}{\pi} \int_0^{\pi/2} (x_0^2 \cos^2 \theta + y_0^2 \sin^2 \theta)^{-1/2} d\theta.
\end{aligned} \tag{2.1}
$$

The discovery of this representation of the limit by an elliptic integral (which does not appear in Lagrange's work) was recorded by Gauss in his diary [7, p. 542] on May 30, 1799, although only for the case $x_0 = 1$, $y_0 = \sqrt{2}$, when the integral equals the quarter-perimeter of Bernoulli's lemniscate. In a memoir of 1818 on the motion of planets, he gave a proof of the general case [6] by substituting

$$
x_0 \csc \theta = x_1 \csc \phi + (x_0 - x_1) \sin \phi \tag{2.2}
$$

in order to show that the integral is unchanged if x_0, y_0 are replaced by x_1, y_1 and eventually by x_n, y_n. Since x_n and y_n approach a common limit, the integrand approaches a constant. A fuller discussion is given by John Todd [13] in a paper which stimulated the present investigation.

In an unpreserved letter to Pfaff in 1800, Gauss suggested taking y_{n+1} to be $(x_{n+1} y_n)^{1/2}$ instead of $(x_n y_n)^{1/2}$. The common limit of x_n and y_n, say $L_{14}(x_0, y_0)$, was apparently known to Gauss and was promptly determined by Pfaff:

$$
\begin{aligned}
&x_{n+1} = \frac{x_n + y_n}{2}, \qquad y_{n+1} = \left(\frac{x_n + y_n}{2} y_n\right)^{1/2}, \qquad n = 0, 1, 2, \cdots, \\
&\frac{1}{L_{14}(x_0, y_0)} = (y_0^2 - x_0^2)^{-1/2} \arccos \frac{x_0}{y_0}, \qquad 0 \leqq x_0 < y_0, \\
&\qquad\qquad = (x_0^2 - y_0^2)^{-1/2} \operatorname{arccosh} \frac{x_0}{y_0}, \qquad 0 < y_0 < x_0.
\end{aligned} \tag{2.3}
$$

Pfaff's reply [7, pp. 234, 284] to Gauss was still unpublished in 1880 when Borchardt rather laboriously rediscovered this algorithm which bears his name [1], [13], [14], [15]. (He died a few months later after twenty-five years as the editor of Crelle's Journal.) A quick proof proceeds by noting that the ratio $r_n = x_n/y_n$ satisfies $r_{n+1} = [\frac{1}{2}(1+r_n)]^{1/2}$. Putting $r_n = \cos(2^{-n}\theta)$ if $x_0 < y_0$, we find that $2^n \arccos(x_n/y_n)$ is independent of n. Now $2^n(y_n^2 - x_n^2)^{1/2}$ also is independent of n. Since x_n and y_n are easily shown to have a common limit, the limit of the invariant

$$\frac{1}{(y_n^2 - x_n^2)^{1/2}} \arccos \frac{x_n}{y_n} = \frac{1}{(y_n^2 - x_n^2)^{1/2}} \arcsin \frac{(y_n^2 - x_n^2)^{1/2}}{y_n}$$

is $1/\lim y_n$, which must therefore equal the value of the invariant for $n = 0$.

The elementary function $1/L_{14}$ is a degenerate case of an elliptic integral and has an integral representation like that of $1/L_{12}$. This remark suggests that the substitution (2.2) might be used for a unified proof of Gauss' and Borchardt's algorithms. With a different choice of integration variable we shall give such a unified proof of not just these two but a full dozen algorithms, among them new algorithms for computing a logarithm and an inverse lemniscatic sine. The one for a logarithm, which has an alternative proof without integrals, is

$$(2.4)\quad \begin{gathered} x_{n+1} = \left(x_n \frac{x_n + y_n}{2}\right)^{1/2}, \quad y_{n+1} = \left(y_n \frac{x_n + y_n}{2}\right)^{1/2}, \quad n = 0, 1, 2, \cdots, \\ [L_{34}(x_0, y_0)]^2 = \frac{x_0^2 - y_0^2}{2 \log (x_0/y_0)}, \qquad x_0 \neq y_0, \end{gathered}$$

where $x_0 > 0$, $y_0 > 0$, and $L_{34}(x_0, y_0)$ is the common limit of x_n and y_n.

The other is the lemniscatic twin of (2.3) and might have pleased Gauss, who developed the theory of lemniscate functions in 1797–98. We define

$$(2.5)\quad \operatorname{arcsl} x = \int_0^x (1 - t^4)^{-1/2} dt, \qquad x^2 \leqq 1,$$

$$(2.6)\quad \operatorname{arcslh} x = \int_0^x (1 + t^4)^{-1/2} dt,$$

where arcsl x is called the **arc lemniscatic sine** of x, i.e., the length of arc of the lemniscate from the origin to the point with radial coordinate x. The notation arcslh x is suggested by the similarity between (2.6) and the integral representation of arcsinh x, but the function does not have as simple a geometric interpretation. The algorithm which rectifies a lemniscatic arc is

$$x_{n+1} = \frac{x_n + y_n}{2}, \qquad y_{n+1} = \left(x_n \frac{x_n + y_n}{2}\right)^{1/2}, \qquad n = 0, 1, 2, \cdots,$$

$$(2.7)\quad [L_{13}(x_0, y_0)]^{-1/2} = (x_0^2 - y_0^2)^{-1/4} \operatorname{arcsl}\left(1 - \frac{y_0^2}{x_0^2}\right)^{1/4}, \qquad 0 \leqq y_0 < x_0,$$

$$= (y_0^2 - x_0^2)^{-1/4} \operatorname{arcslh}\left(\frac{y_0^2}{x_0^2} - 1\right)^{1/4}, \qquad 0 < x_0 < y_0,$$

where $x_n, y_n \to L_{13}(x_0, y_0)$. The real period of the lemniscate functions can be computed iteratively either by (2.1), as Gauss discovered in 1799, or by putting $x_0 = 1$, $y_0 = 0$ in (2.7). From the proof of (2.7) we shall deduce also Fagnano's duplication theorem.

3. Proof of twelve algorithms. In Lemma 1 we define sixteen pairs of sequences $\{x_n\}$, $\{y_n\}$ and show that the two sequences of each pair have a common limit. The limit is then evaluated for each case in Theorem 1. In the twelve nontrivial cases, including Gauss' and Borchardt's algorithms, the limit is expressed in terms of a possibly degenerate elliptic integral.

LEMMA 1. *Let x_0 and y_0 be positive numbers. For any fixed choice of the indices i and j from among the numbers* 1, 2, 3, 4 *we define*

$$(3.1)\qquad x_{n+1} = f_i(x_n, y_n), \qquad y_{n+1} = f_j(x_n, y_n), \qquad n = 0, 1, 2, \cdots,$$

where

$$(3.2)\qquad \begin{aligned} f_1(x, y) &= \frac{x + y}{2}, & f_2(x, y) &= (xy)^{1/2}, \\ f_3(x, y) &= \left(x\,\frac{x + y}{2}\right)^{1/2}, & f_4(x, y) &= \left(\frac{x + y}{2}\,y\right)^{1/2}, \end{aligned}$$

all square roots taken positive. Then the sequences $\{x_n\}$ and $\{y_n\}$ have a common limit which we denote by $L_{ij}(x_0, y_0)$:

$$(3.3)\qquad \lim_{n\to\infty} x_n = \lim_{n\to\infty} y_n = L_{ij}(x_0, y_0).$$

Proof. Let x and y be positive and let f_i stand for $f_i(x, y)$. It is trivial to verify that exactly one of the following statements is true according as $x - y$ is negative, zero, or positive:

$$(3.4)\qquad \begin{aligned} &x < f_3 < f_2 < f_1 < f_4 < y, \\ &x = f_1 = f_2 = f_3 = f_4 = y, \\ &y < f_4 < f_2 < f_1 < f_3 < x. \end{aligned}$$

Let I_n be the closed interval of the real line with endpoints x_n^2 and y_n^2, possibly degenerating to a single point if $x_n = y_n$. Then (3.1) and (3.4) imply $I_n \supset I_{n+1}$. Moreover, for any choice of i and j we have

$$(3.5)\qquad | x_{n+1}^2 - y_{n+1}^2 | \leqq | f_3^2(x_n, y_n) - f_4^2(x_n, y_n) | = \frac{| x_n^2 - y_n^2 |}{2}.$$

Thus $I_0, I_1, I_2, \cdots$ are nested intervals with lengths tending to zero, and their intersection is a single point which is the common limit of the sequences $\{x_n^2\}$ and $\{y_n^2\}$. Since x_n and y_n are positive, the sequences $\{x_n\}$ and $\{y_n\}$ also have a common limit.

We shall now express the limits in terms of integrals of the form

$$R(a; b, b'; x^2, y^2) = \frac{1}{B(a, a')} \int_0^\infty t^{a'-1}(t + x^2)^{-b}(t + y^2)^{-b'} dt, \tag{3.6}$$

where B denotes the beta function, a' is defined by

$$a + a' = b + b', \tag{3.7}$$

and we assume $\operatorname{Re} a > 0$ and $\operatorname{Re} a' > 0$. The R-function [2], [4] is homogeneous of degree $-a$ in the arguments x^2 and y^2 and has the value unity if $x^2 = y^2 = 1$. It has the symmetry property

$$R(a; b, b'; x^2, y^2) = R(a; b', b; y^2, x^2) \tag{3.8}$$

and is related to the Gauss hypergeometric function by

$$R(a; b, b'; x^2, y^2) = y^{-2a} {}_2F_1\left(a, b; b + b'; 1 - \frac{x^2}{y^2}\right). \tag{3.9}$$

The content of the following theorem could in principle be extracted from known quadratic transformations of the hypergeometric function:

THEOREM 1. *Let x_0 and y_0 be positive numbers and define sequences $\{x_n\}$ and $\{y_n\}$ as in Lemma 1 for each choice of i and j. The common limit of the two sequences is specified by the entry in the ith row and jth column of the following table:*

$i \backslash j$	1	2	4	3
1	f_1	$(\frac{1}{2}; \frac{1}{2}, \frac{1}{2})$	$(\frac{1}{2}; \frac{1}{2}, 1)$	$(\frac{1}{4}; \frac{3}{4}, \frac{1}{2})$
2	$(\frac{1}{2}; \frac{1}{2}, \frac{1}{2})$	f_2	$(1; \frac{1}{2}, 1)$	$(1; \frac{3}{4}, \frac{1}{2})$
3	$(\frac{1}{2}; 1, \frac{1}{2})$	$(1; 1, \frac{1}{2})$	$(1; 1, 1)$	f_3
4	$(\frac{1}{4}; \frac{1}{2}, \frac{3}{4})$	$(1; \frac{1}{2}, \frac{3}{4})$	f_4	$(1; 1, 1)$

If $i = j$ the entry f_i means that $x_n = y_n = f_i(x_0, y_0)$ for $n = 1, 2, 3 \cdots$. If $i \neq j$ the entry $(a; b, b')$ means that $R(a; b, b'; x_n^2, y_n^2)$ is independent of n for $n = 0, 1, 2, \cdots$, and

$$L_{ij}(x_0, y_0) = \left[R(a; b, b'; x_0^2, y_0^2)\right]^{-1/2a}. \tag{3.10}$$

Proof. We may assume $i \neq j$, for $i = j$ implies $x_1 = y_1 = x_2 = y_2 = \cdots$. Letting f_i stand for $f_i(x, y)$, we make in (3.6) the substitution

$$t = \frac{s(s + f_2^2)}{s + f_1^2},$$

$$\frac{dt}{ds} = \frac{s^2 + 2sf_1^2 + f_1^2 f_2^2}{(s + f_1^2)^2} = \frac{(s + f_3^2)(s + f_4^2)}{(s + f_1^2)^2}, \tag{3.11}$$

$$t + x^2 = \frac{(s + f_3^2)^2}{s + f_1^2}, \qquad t + y^2 = \frac{(s + f_4^2)^2}{s + f_1^2},$$

with the result that

$$R(a; b, b'; x^2, y^2) = \frac{1}{B(a, a')} \int_0^\infty s^{a'-1} (s + f_1^2)^{a-1} (s + f_2^2)^{a'-1} (s + f_3^2)^{1-2b} (s + f_4^2)^{1-2b'} ds. \tag{3.12}$$

Given i and j we now choose two parameters so that f_k disappears from the integrand if $k \neq i$ or j. Since all cases are similar, we shall carry through only the case $i = 1$ and $j = 3$. Choosing $a' = 1$ and $b' = 1/2$ and using (3.6) and (3.7), we find

$$R(a; a + \tfrac{1}{2}, \tfrac{1}{2}; x^2, y^2) = R(a; 1 - a, 2a; f_1^2, f_3^2). \tag{3.13}$$

The first parameter and the sum of the second and third parameters are the same on both sides as a consequence of (3.12). Thus there is exactly one value of a, viz., $a = 1/4$, for which each parameter is the same on both sides. Defining

$$R_A(x^2, y^2) = R(\tfrac{1}{4}; \tfrac{3}{4}, \tfrac{1}{2}; x^2, y^2), \tag{3.14}$$

we have

$$R_A(x^2, y^2) = R_A(f_1^2, f_3^2). \tag{3.15}$$

If sequences $\{x_n\}$ and $\{y_n\}$ are defined as in Lemma 1 with $i = 1$ and $j = 3$, it follows that $R_A(x_n^2, y_n^2)$ is independent of n. Letting L stand for $L_{13}(x_0, y_0)$ and noting that R is continuous for positive values of its arguments [2], [4], we have

$$R_A(x_0^2, y_0^2) = \lim_{n \to \infty} R_A(x_n^2, y_n^2) = R_A(L^2, L^2) = (L^2)^{-1/4} R_A(1, 1) = L^{-1/2}. \tag{3.16}$$

By interchanging x and y, which entails interchanging f_3 and f_4, we find $R_A(y^2, x^2) = R_A(f_1^2, f_4^2)$ and hence, by the symmetry property (3.8),

$$R(\tfrac{1}{4}; \tfrac{1}{2}, \tfrac{3}{4}; x^2, y^2) = R(\tfrac{1}{4}; \tfrac{1}{2}, \tfrac{3}{4}; f_4^2, f_1^2). \tag{3.17}$$

The same procedure as before leads now to

$$L_{41}(x_0, y_0) = \left[R(\tfrac{1}{4}; \tfrac{1}{2}, \tfrac{3}{4}; x_0^2, y_0^2)\right]^{-2}. \tag{3.18}$$

The columns of the table have been arranged so that two results connected in this way by symmetry are found in corresponding positions on opposite sides of the main diagonal.

Although x_0 and y_0 were assumed positive, it happens in some cases that the integral in (3.6) exists if $x=0$ and $y\neq 0$. In such cases (3.10) is still valid if $x_0=0$, for both sides of (3.10) are then continuous at $x_0=0$. Similar remarks apply to y_0.

4. Discussion. We indicate first how to get the algorithms in Section 2 from those of Theorem 1. The representation (3.6) of $R(\frac{1}{2}; \frac{1}{2}, \frac{1}{2}; x_0^2, y_0^2)$ is changed into the integral in Gauss' algorithm (2.1) by putting $t=x_0^2 \cot^2\theta$. Moreover, if we square both sides of Gauss' substitution (2.2) and put $\cot^2\theta=t/x_0^2$ and $\cot^2\phi=s/x_1^2$, where $x_1=\frac{1}{2}(x_0+y_0)$, we find $t=s(s+x_0y_0)/(s+x_1^2)$, which is equivalent to the first equation of (3.11).

In the representation (3.6) of $R(\frac{1}{2}; \frac{1}{2}, 1; x_0^2, y_0^2)$ we put $(t+x_0^2)/(t+y_0^2)$ equal to $\cos^2\theta$ or $\cosh^2\theta$ according as $x_0<y_0$ or $x_0>y_0$ in order to obtain Borchardt's algorithm (2.3). The logarithmic algorithm (2.4) comes similarly from $R(1; 1, 1; x_0^2, y_0^2)$ by putting $(t+x_0^2)/(t+y_0^2)=e^\theta$.

In the representation (3.6) of $R_A(x_0^2, y_0^2)$ we put $(t+y_0^2)/(t+x_0^2)$ equal to $1-T^4$ or $1+T^4$ according as $x_0>y_0$ or $x_0<y_0$ to obtain the lemniscatic algorithm (2.7). In particular we find from (2.5) and (2.6) that

$$\operatorname{arcsl} x = xR_A(1, 1-x^4), \qquad \operatorname{arcslh} x = xR_A(1, 1+x^4). \tag{4.1}$$

The invariance of $R_A(x_n^2, y_n^2)$ under $n\to n+1$ (see Theorem 1) implies

$$\begin{aligned} \operatorname{arcsl} x &= \sqrt{2}\operatorname{arcslh} y, \quad x = \sqrt{2}\,y(1+y^4)^{-1/2}, \qquad y^2 \leqq 1, \\ \operatorname{arcslh} y &= \sqrt{2}\operatorname{arcsl} z, \quad y = \sqrt{2}\,z(1-z^4)^{-1/2}, \qquad z^2 < 1. \end{aligned} \tag{4.2}$$

The arcsl function is transformed into arcslh because the recurrence relations in (2.7) give $x_{n+1}<y_{n+1}$ if $x_n>y_n$. In proving (4.2) we recall that R_A is homogeneous of degree $-\frac{1}{4}$ in its arguments. For example,

$$\begin{aligned} \operatorname{arcsl} x &= xR_A(1, 1-x^4) = xR_A(x_1^2, x_1) \\ &= \frac{x}{\sqrt{x_1}} R_A\left(1, \frac{1}{x_1}\right) = \sqrt{2}\,yR_A(1, 1+y^4) = \sqrt{2}\operatorname{arcslh} y, \end{aligned} \tag{4.3}$$

where

$$x_1 = \frac{1+(1-x^4)^{1/2}}{2} = \frac{1}{1+y^4}\,. \tag{4.4}$$

Combining the two parts of (4.2), we see that invariance under $n\to n+2$ implies

$$\operatorname{arcsl} x = 2\operatorname{arcsl} z, \quad x = \frac{2z(1-z^4)^{1/2}}{1+z^4}, \qquad z^2 \leqq \sqrt{2}-1. \tag{4.5}$$

This is the duplication theorem for an arc of Bernoulli's lemniscate [11], [16], [9]. Since the relation between x and z involves only rational operations and a square root, the duplication can be effected, as Fagnano observed, with ruler and compass. The restriction on z ensures that $y^2 \leqq 1$ and hence $x^2 \leqq 1$.

The remaining algorithms of Theorem 1 could be deduced from the four already discussed by using the symmetry and homogeneity of the R-function together with its Euler transformation [2, Equation (2.8)]. Only Gauss' algorithm enjoys quadratic convergence; it is easy to see from (3.2) that $|x_n - y_n|$ ultimately decreases by a factor 4 in each step of iteration in the other cases, except for the logarithmic case with factor 2.

An elementary proof of (2.4) can be generalized as follows for any $p>1$: If x_0 and y_0 are positive and unequal, we define

$$x_{n+1} = \left[\frac{x_n(x_n^p - y_n^p)}{p(x_n - y_n)}\right]^{1/p}, \qquad y_{n+1} = \left[\frac{y_n(x_n^p - y_n^p)}{p(x_n - y_n)}\right]^{1/p}, \quad n = 0, 1, 2, \cdots. \tag{4.6}$$

It is easily verified that x_n and y_n approach a common limit, say L, and that $p^n \log(x_n/y_n)$ and $p^n(x_n^p - y_n^p)$ are independent of n. The limit of the invariant $(x_n^p - y_n^p)/\log(x_n/y_n)$ is $p \lim x_n^p$, and hence pL^p must equal the value of the invariant for $n=0$:

$$L^p = \frac{x_0^p - y_0^p}{\log x_0^p - \log y_0^p}. \tag{4.7}$$

If we replace y_m by $(x_m y_m)^{1/2}$ for $m=0, 1, 2, \cdots$, in Borchardt's algorithm (2.3), we find

$$\begin{gathered} x_{n+1} = \frac{x_n + (x_n y_n)^{1/2}}{2}, \qquad y_{n+1} = (x_n y_n)^{1/2}, \quad n = 0, 1, 2, \cdots, \\ \frac{1}{\lim x_n} = \frac{1}{\lim y_n} = R(\tfrac{1}{2}; \tfrac{1}{2}, 1; x_0^2, x_0 y_0). \end{gathered} \tag{4.8}$$

This algorithm was given by both Pfaff [7, p. 235] and Borchardt [1]. Other variants can be obtained by similar substitutions.

For each of the twelve algorithms in Theorem 1, the limit (3.10) is a homogeneous mean value of x_0 and y_0 of the type discussed by Tobey [12]. Gauss' and Borchardt's algorithms are special cases of three-dimensional algorithms [3], [5].

Work performed in the Ames Laboratory of the U. S. Atomic Energy Commission. The author thanks Professor John Todd for the initial stimulus provided by an unpublished manuscript [13] as well as for several references and comments.

References

1. C. W. Borchardt, Sur deux algorithmes analogues à celui de la moyenne arithmético-géométrique de deux éléments, In Memoriam Dominici Chelini, Collect. Math., [etc.], L. Cremona,

ed., U. Hoepli, Milan, 1881, pp. 206–212; reprinted in C. W. Borchardt, Gesammelte Werke, Berlin, 1888, pp. 455–462.

2. B. C. Carlson, Lauricella's hypergeometric function F_D, J. Math. Anal. Appl., 7 (1963) 452–470. MR 28 #258.

3. ———, On computing elliptic integrals and functions, J. Math. and Phys., 44 (1965) 36–51. MR 30 #5470.

4. ———, A connection between elementary functions and higher transcendental functions, SIAM J. Appl. Math., 17 (1969) 116–148.

5. ———, Hidden symmetries of special functions, SIAM Rev., 12 (1970) 332–345.

6. C. F. Gauss, Werke, vol. 3, Teubner, Leipzig, 1876, pp. 352–355.

7. ———, Werke, vol. 10, part 1, Teubner, Leipzig, 1917.

8. J.-L. Lagrange, Oeuvres, vol. 2, Gauthier-Villars, Paris, 1868, pp. 253–312, esp. pp. 267, 272.

9. A. I. Markushevich, The Remarkable Sine Functions, American Elsevier, New York, 1966.

10. L. Schlesinger, Über Gauss' Jugendarbeiten zum arithmetisch-geometrischen Mittel, Jber. Deutsch. Math.-Verein., 20 (1911) 396–403.

11. C. L. Siegel, Topics in Complex Function Theory, vol. 1, Wiley-Interscience. New York, 1969, pp. 1–7.

12. M. D. Tobey, A two-parameter homogeneous mean value, Proc. Amer. Math. Soc., 18 (1967) 9–14. MR 35 #1742.

13. John Todd, The lemniscate constants, to appear.

14. F. G. Tricomi, Sull'algoritmo iterativo del Borchardt e su di una sua generalizzazione, Rend. Circ. Mat. Palermo, (2) 14 (1965) 85–94.

15. ———, Lectures on the use of special functions by calculations with electronic computers, Lecture Series No. 47, The Institute for Fluid Dynamics and Applied Mathematics, University of Maryland, College Park, 1966.

16. G. N. Watson, The marquis and the land-agent; a tale of the eighteenth century, Math. Gaz., 17 (1933) 5–17.

17. D. H. Lehmer, On the compounding of certain means, J. Math. Anal. Appl., to appear; Notices Amer. Math. Soc., 17 (1970) 634–635.

18. P. G. Stäckel and W. Ahrens, eds., Der Briefwechsel zwischen C. G. J. Jacobi und P. H. von Fuss über die Herausgabe der Werke Leonhard Eulers, Teubner, Leipzig, 1908, pp. 23, 31.

A NOTE ON THE UNIVERSAL COVERING SPACE OF A SURFACE

G. W. KNUTSON, University of Nebraska

The purpose of this note is to prove geometrically that the universal covering space of a surface, other than the 2-sphere or the projective plane, is the plane. This will be done by constructing the covering of the double torus by the plane and extending this construction to the remaining surfaces.

Let X and $X^\#$ be topological spaces. $X^\#$ is called a **covering space** of X if there is a continuous map p from $X^\#$ onto X such that for each point x of X there is an open neighborhood U of x such that $p^{-1}(U)$ is the disjoint union of open sets U_α and each $p|U_\alpha$ is a homeomorphism of U_α onto U. If $X^\#$ is simply connected, i.e., $X^\#$ is connected, and each map on the boundary of the unit disk into $X^\#$ has a continuous extension over the entire disk, then $X^\#$ is called a **universal**

covering space of X. It is not hard to see that if $X^\#$ is the universal covering space of X, then it is unique up to homeomorphism [2, p. 160].

If X is the sphere or the projective plane, then $X^\#$, the universal covering space of X, is the 2-sphere. If X is a torus, then $X^\#$ is the plane [2, p. 147]. We shall, however, look closely at this covering of the torus in order to extend it to the double torus.

We first cut the torus in half, obtaining two annuli. Inside the "front" annulus we find a spine (a spine of a manifold with boundary is a connected set with a complement homeomorphic to the product of the boundary of the manifold with a half-open interval). Clearly the central curve of an annulus is a spine. Since the universal covering space of a simple closed curve is the real line [1], we may extend this covering to a covering of the annulus by the infinite band $(-\infty, \infty)\times[0, 1]$. Now attach a countable collecting of closed bands, say $(-\infty, \infty)\times[n, n+1]$ where n is an integer, obtaining the plane. Since each band covers the annulus, by alternately covering the "front" and "back" annuli of the torus with alternate bands we obtain the universal covering space of the torus.

In the case of the double torus we use exactly the same technique. By cutting the double torus in half we obtain two double annuli. If Y is the front double annulus, then Y has a spine X, where X is homeomorphic to a figure eight. (See Figure 1.)

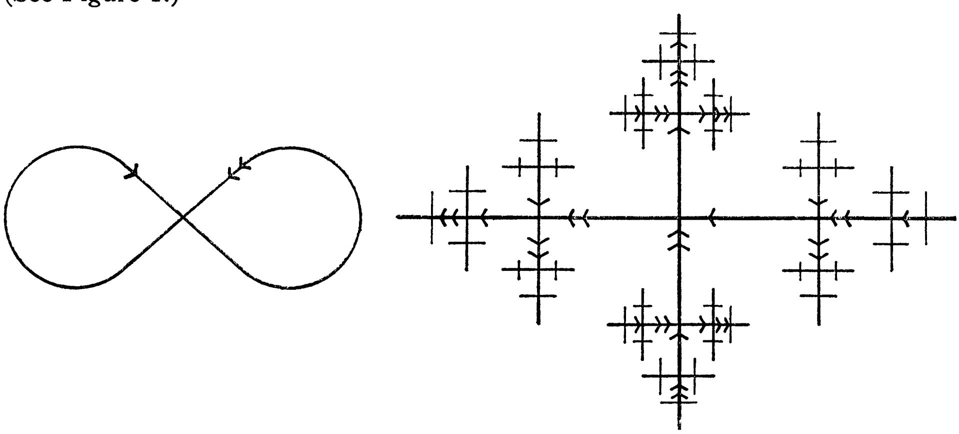


FIG. 1

The universal covering space $X^\#$ of X is pictured in figure 1. To construct $X^\#$, start with an open segment of unit length. At the midpoint of this segment, erect an open interval of length 2/3 so that the two intervals are perpendicular bisectors of each other. At the midpoints of the four resulting half-open segments, erect perpendicular bisectors of length 2/3 of that of the segments being bisected. Continue this process on each of the resulting half-open segments. Inductively we obtain $X^\#$. To see that $X^\#$ is a covering space of X, orient the 1-spheres of X and the segments of $X^\#$ as indicated in figure 1. Notice that the

direction on each horizontal and vertical line remains constant, but that the segments are alternately indexed. This insures that at each intersection point of $X^\#$ the four line segments have the same local orientation as the segments of the figure eight. The covering map takes the segments with a single index onto the left circle of X and the segments with a double index onto the right circle of X in an orientation preserving manner.

We now need to construct a space $Y^\#$ which has $X^\#$ as a spine and is the universal covering space of Y. Consider a closed segment S of $X^\#$ of length β. Replace S with a symmetric trapezoid with S as its center of symmetry, having height β and bases of length $\beta/10$ and $\beta/20$. This is the same as replacing each of the half-open segments of $X^\#$ by a cone less its vertex. See figure 2. By construction, it follows that $Y^\#$ is an open disk with a countable number of open segments on the boundary.

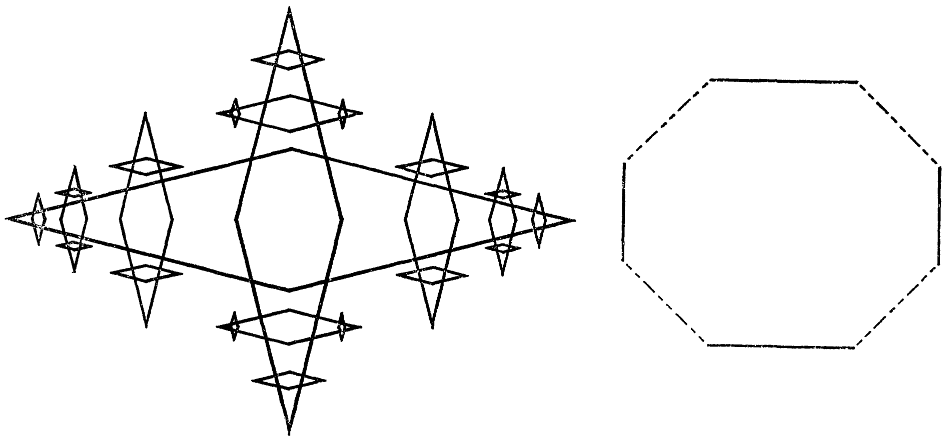


FIG. 2

Since we shall need a more convenient form of $Y^\#$, we consider the following homeomorphic image of $Y^\#$. Start with an open octagon. Along the edges parallel to the axes add open segments. On each of the remaining edges we add a pair of open segments. We continue to add pairs of open segments in each of the resulting spaces. This octagonal form of $Y^\#$ is pictured in figure 2.

To cover Y with $Y^\#$, we need only extend the map from $X^\#$ to X. This can be accomplished by extending the map of each segment to the corresponding trapezoid by preserving the product structure. The orientation given to $X^\#$ guarantees that this map can be made to agree on each trapezoid and so is an extension. Thus $Y^\#$ is a covering space of Y.

To construct the universal covering space of the double torus, we connect a countable collection of the octagonal form of $Y^\#$. Consider the above octagon as a generating octagon. Along each segment of this octagon attach another octagon by identifying one of the boundary components. To each segment of the generating octagon we attach an octagon. Each of these octagons is in the

second generation. Along each segment of the resulting space, attach another octagon. Each of these octagons is in the third generation. Continue this process. The resulting figure is clearly an open disk.

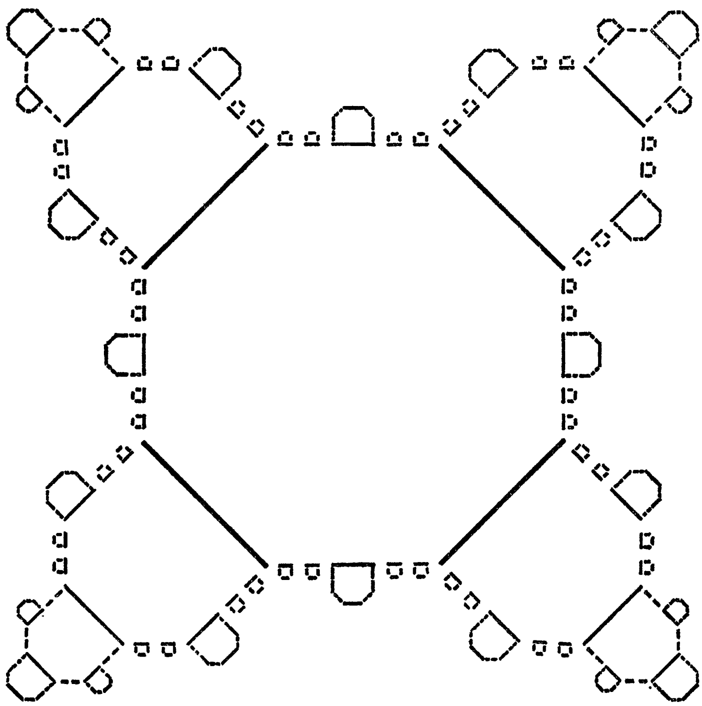


FIG. 3

To see that the open disk is a covering space of the double torus, we must establish a covering map. From the second step we have a map from the generating octagon onto the double annulus that comprises the "front" of the double torus. For any second generation octagon, we have the map defined on one of the segments on its boundary. Thus there is no difficulty in extending the map, by means of the covering map of $Y^{\#}$ onto Y, to a map of the octagon onto the "back" of the double torus. Thus we may extend the map over the second generation. In the same way we extend the map over the third and subsequent generations. In general the $2n$th generation is mapped onto the "back" and the $2n+1$st generation is mapped onto the "front" of the double torus. Clearly the resulting map is a covering map. Thus the plane is the universal covering space of the double torus.

For a sphere with n handles, we repeat the same argument. Instead of an open octagon we obtain an open $4n$-gon in the second step. It is interesting to note, however, that the $4n$-gon is homeomorphic to the generating octagon of the preceding argument. For a sphere with n crosscaps we need the following facts:

FACT I: If $X^\#$ is a covering space of X and $Y^\#$ is a universal covering space of $X^\#$, then $Y^\#$ is a universal covering space of $X^\#$, [2, p. 161].

FACT II: A sphere with $2n$-handles is a 2-fold covering of a sphere with $2n+1$ crosscaps and a sphere with $2n-1$ handles is a 2-fold covering of a sphere with $2n$ crosscaps, [1, p. 118].

Combining the above we have established the following theorem:

THEOREM. *If S is a surface other than a 2-sphere or a projective plane, S has the plane as a universal covering space.*

References

1. M. Greenberg, Lectures on Algebraic Topology, Benjamin, New York, 1967.

2. W. Massey, Algebraic Topology: An Introduction, Harcourt, Brace and World, New York, 1967.

SOME CONSEQUENCES OF THE UNIVERSAL CHORD THEOREM

J. T. ROSENBAUM, University of Pittsburgh

The Universal Chord Theorem (UCT) states that the values $\alpha=1$, $1/2$, $1/3, \cdots$ are precisely those for which the following statement is true:

If f is a continuous, real valued function on the closed unit interval such that $f(0)=f(1)$, then the graph of f has a chord of length α. (The word **chord** *will mean horizontal chord.)*

This theorem was discovered by P. Levy [3] and generalized by H. Hopf [2]. Both dealt with a plane continuum (a bounded, closed, connected set) K, of which the graph of a function f of the type we consider is a special case. We present elementary proofs of two generalizations of this special case, and give two applications. The first application will shed light on the Universal Chord Theorem, while the second is a converse of the local invertibility theorem for an analytic function near a point where its derivative doesn't vanish.

Levy proved that if K is a plane continuum, $S(K)$ is the set of chordal lengths for K, and c is a number in $S(K)$, then c/n belongs to $S(K)$ for each natural number n. Also, if $0<\alpha<1$ and α is not of the form $1/n$, then there exists a K such that 1 is in $S(K)$ but α isn't.

Hopf proved that a set M is $S(K)$ for some K if and only if M^* (where * denotes complement with respect to the positive reals) is open, non-empty, and closed under addition.

Assuming Hopf's result, suppose c/n does not belong to $S(K)$. Then c/n is a member of the additive set $S^*(K)$, so $c/n+ \cdots +c/n$ (n terms) belongs to $S^*(K)$, i.e., c is not a member of $S(K)$. Conversely, if $1/\alpha$ is not integral, let n be the natural number determined by $n<1/\alpha<n+1$, and define M by M^*

$=\bigcup_1^\infty (k/(n+1), k/n)$, so M^* is open and non-empty. We easily see that M^* is additive, so the K supplied by Hopf's work serves for the other half of Levy's theorem, since

$$1 \in M \Rightarrow 1 \in S(K) \quad \text{and} \quad \alpha \notin M \Rightarrow \alpha \notin S(K).$$

Levy's result implies half of the UCT since the graph of f is a plane continuum. The other half of the UCT is established by a counterexample. (Cf. remark 1 below.)

Proof of half of the UCT. Given f and the natural number n, if $f(t+1/n)-f(t)$ were nowhere 0 on $[0, (n-1)/n]$, then by the intermediate value theorem it would be either strictly positive or strictly negative. This is impossiple since

$$0 = f(1) - f(0) = \sum_0^{n-1}\left[f\left(\frac{k}{n}+\frac{1}{n}\right)-f\left(\frac{k}{n}\right)\right].$$

GENERALIZATION 1. *For each f and n, the graph of f has at least n chords whose lengths are multiples of $1/n$.*

Proof 1. Levy proved that if $0<\alpha<1$ and f does not have a chord of length α then it has two chords of length $1-\alpha$. For a simple proof of this, see Boas [1]. Using this result, let $1 \leqq k \leqq n-1$. If f doesn't have a chord of length k/n, it has two chords of length $(n-k)/n$; if it doesn't have one of length $(n-k)/n$, it has two of length k/n. Thus f has at least as many chords whose lengths are in the set $\{k/n, (n-k)/n\}$ as this set has members. Thus since f always has a chord of length n/n, we can always find the required number of chords.

Proof 2. If the statement holds for n, fix f. The UCT supplies us with a chord of length $1/(n+1)$ as shown in Fig. 1.

Remove the portion of the graph between the endpoints P_1 and P_2 of the chord and slide the pieces together in such a way that P_1 and P_2 become coincident, say at P, and that the resulting function f_0 has the interval $[0, n/(n+1)]$ as its domain. This can be done more formally by composing f with an appropriate Heaviside function. The induction hypothesis, with a scale change, can be applied to obtain n chords whose lengths are multiples of $(1/n)[n/(n+1)]$ $=1/(n+1)$. Each of these gives rise to a chord on the graph of f that is either of the same length or is $1/(n+1)$ longer, according to whether the f_0 chord fails or does not fail to straddle the point P. These n chords together with the original one furnished by the UCT form the required number, to complete the induction.

GENERALIZATION 2. *If g is continuous and strictly monotone on $[0, 1]$ and $L=[g(1)-g(0)]$, then for any f there exist n pairs $\{T_j, T_j'\}$ in the unit interval such that $f(T_j)=f(T_j')$ and such that $|g(T_j)-g(T_j')|$ is a multiple of $|L/n|$.*

Proof. Apply the first generalization to

$$F = f\circ\left[\frac{1}{L}(g-g(0))\right]^{-1}$$

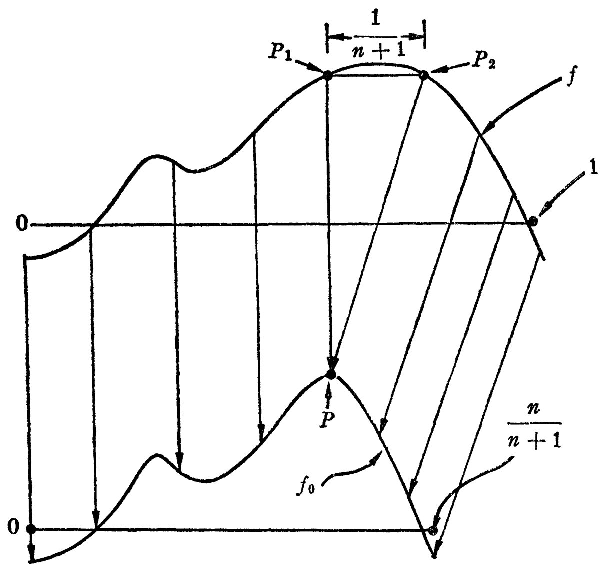


FIG. 1

to get n pairs $\{t_j, t_j'\}$, and let

$$T_j = \left[\frac{1}{L}(g - g(0))\right]^{-1}(t_j), \qquad T_j' = \left[\frac{1}{L}(g - g(0))\right]^{-1}(t_j').$$

APPLICATION 1. *A closed planar curve which monotonically turns about a point n times must intersect itself at least $n-1$ times.*

Proof. We are supposing that the curve has the form

$$w_0 + f(t) \exp i\theta(t),$$

where θ maps the unit interval strictly monotonically and continuously onto an interval of length $2\pi n$. Generalization 2 is immediately applicable with $\theta = g$. Since one of the pairs $\{T_j, T_j'\}$ may be $\{0, 1\}$ and since we don't count the initial-terminal coincidence as an intersection, we are thus left with $n-1$ intersections (counting multiplicity).

APPLICATION 2. *Suppose $\phi(z) = w_0 + a(z-z_0)^n\psi(z)$ $(a \neq 0,\ \psi(z_0) = 1,\ |z-z_0| > R > 0)$, where ψ is analytic at z_0. We lose no generality (R can be made smaller) in supposing that $|\psi(z) - 1| < 1$. Thus, we can write*

$$\psi(z) = |\psi(z)| \exp i\theta_0(z),$$

where $\theta_0(z) = \tan^{-1}(\operatorname{Im}\psi(z)/\operatorname{Re}\psi(z))$. Thus

$$\phi(z_0 + r \exp 2\pi i t) = w_0 + f(t) \exp i\theta(t),$$

where $f(t) = |ar^n\psi(z_0 + r\exp 2\pi i t)|$ *and*

$$\theta(t) = 2\pi n t + \theta_0(z_0 + r\exp 2\pi i t).$$

The conclusion of application 1 is what is needed now while the hypotheses are readily verified except, perhaps, for monotonicity. A simple computation shows that

$$\frac{d}{dt}\theta_0(z_0 + r\exp 2\pi i t)$$

tends to 0 uniformly as r tends to 0 so, for r sufficiently small, it is dominated by $(d/dt)2\pi n t$ insuring that θ' is strictly positive.

REMARK 1. The idea of application 1 can be used to supply a proof of the "missing" half of the UCT as follows: if α is a positive number whose reciprocal is not integral then it is clear that we can monotonically turn about a point $1/\alpha$ times and return to the original radial distance without intersections. Figure 2 illustrates this for the case $\alpha = 4/11$. If $r = r(\theta)$ corresponds to such a curve, then

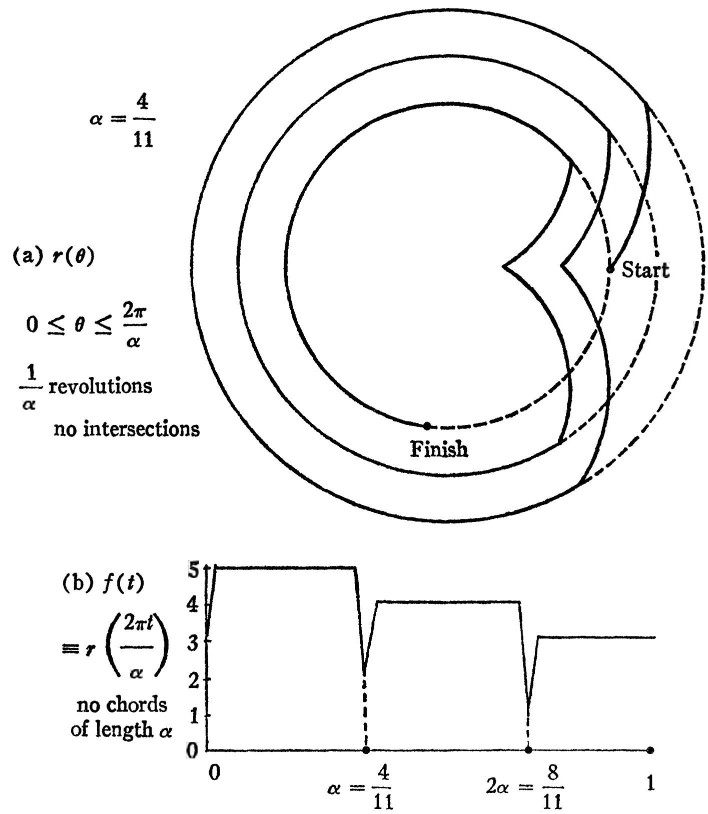


FIG. 2

we set $f(t)=r(2\pi t/\alpha)$. Then $f(t+\alpha)=f(t)$ must fail for each t for otherwise $r(\theta)=r(\theta+2\pi)$, where $\theta=2\pi t/\alpha$. The f corresponding to the $r(\theta)$ of figure 2a is shown in figure 2b.

REMARK 2. Application 1 was proved by Hopf in the absence of the monotonicity requirement but we have not been able to adapt our arguments to this case.

References

1. R. P. Boas, A Primer of Real Functions, Wiley, New York, 1960.
2. H. Hopf, Über die Sehnen ebener Kontinuen und die Schleifen geschlossener Wege, Comment. Math. Helv., 9 (1936–7) 303–319.
3. P. Levy, Sur une Généralisation du Théorème de Rolle, C. R. Acad. Sci., Paris, 198 (1934) 424–425.

MATHEMATICAL NOTES

EDITED BY DAVID DRASIN

Manuscripts for this Department should be sent to Robert Gilmer, Department of Mathematics, Florida State University, Tallahassee, FL 32306.

NONSPHERICAL BODIES WITH CONSTANT *HA*-MEASUREMENTS EXIST

JOSEPH ZAKS, University of Washington and Wayne State University

The purpose of this note is to answer a question raised recently by V. Klee in [1].

A **body** B in the Euclidean 3-space E^3 is a nonempty bounded subset of E^3 which is the closure of its interior. The ***HA*-measurement** of B **relative to a plane** P is the maximum area $A(B\cap P')$ of $B\cap P'$, for P' parallel to P, and it is denoted by $M_P(B)$. The body B is said to have **constant *HA*-measurements** if $M_P(B)=M_Q(B)$ for all planes P and Q.

The body B is called **spherical** if it is the union of concentric spheres.

Our negative answer, as stated in the title, is to the following question (see [1]):

"Is a body spherical if it has constant HA-measurements?"

Some additional properties were proposed in [1] for incorporation into the definition of "body." These were: B is connected, the boundary of B is connected, B is contractible, B is starshaped, and B is convex. Our nonspherical body with constant HA-measurements is centrally symmetric with connected boundary, or else contractible with connected boundary, or it can be chosen to be topologically quite complicated. On the other hand, the sphericity of a convex body with constant HA-measurements is known when the body is centrally symmetric, but not otherwise (see [1]).

Let $A(X)$ denote the area of X, and for $0<a<b$ let

$$B_{ab} = \{(x, y, z) \in E^3 \mid a^2 \leqq x^2 + y^2 + z^2 \leqq b^2\}.$$

We need the following elementary lemmas:

LEMMA 1. *If P is a plane with distance c to the origin, then*

$$A(P \cap B_{ab}) = M_P(B_{ab}) = \pi(b^2 - a^2),$$

provided $0 \leqq c \leqq a$.

Proof. Without loss of generality, we may take P as the plane given by $x=c$. Then $P\cap B_{ab}$ is a planar domain bounded by two concentric circles of radii y_1 and y_2 such that $c^2+y_1^2=a^2$ and $c^2+y_2^2=b^2$. As a result

$$A(P \cap B_{ab}) = \pi(y_2^2 - y_1^2) = \pi[(b^2 - c^2) - (a^2 - c^2)] = \pi(b^2 - a^2),$$

and hence it does not depend on c.

If $c>a$, then $P\cap B_{ab}$ is either a circular disk with radius $(b^2-c^2)^{1/2}$, or it is empty; therefore $A(P\cap B_{ab})<\pi(b^2-a^2)$. This implies $M_P(B_{ab})=\pi(b^2-a^2)$ $=A(P\cap B_{ab})$, and the proof is complete.

COROLLARY. *The body B_{ab} has constant HA-measurements equal to $\pi(b^2-a^2)$.*

(See [1]; note that on p. 540, line 20 should be $(r^2-1)^{1/2}$ instead of $(r^3-1)^{1/3}$.)

Let C be defined by $C=B_{ab}-\{(x, y, z) \mid x<0 \text{ and } y^2+z^2<a^2\}$, so C is obtained from B_{ab} by removing a cylindrical plug going from one boundary component of B_{ab} to the other one.

THEOREM. *Every set X with $C\subseteq X\subseteq B_{ab}$ has constant HA-measurements.*

Proof: Let X be such that $C\subseteq X\subseteq B_{ab}$. From the definition of HA-measurements it follows that for all P,

$$M_P(C) \leqq M_P(X) \leqq M_P(B_{ab}) = \pi(b^2 - a^2),$$

where the equality is due to the corollary. Consequently, $M_P(C) \leqq \pi(b^2-a^2)$, for all P.

To show that $M_P(C)\geqq\pi(b^2-a^2)$ for all P, let P be a plane, and let L be a line perpendicular to P passing through the origin. Let

$$S = \{(x, y, z) \mid x \geqq 0 \text{ and } x^2 + y^2 + z^2 = a^2\}.$$

Then S is a closed half-sphere centered at the origin, hence S meets each ray emanating from the origin; in particular, $L\cap S\neq\emptyset$. Let $p\in L\cap S$, and let P' be the plane parallel to P through p. Then $C\cap P'$ is a circular disk with radius $(b^2-a^2)^{1/2}$; therefore $M_P(C)\geqq\pi(b^2-a^2)$.

As a result, $M_P(C)=\pi(b^2-a^2)$, for all P; this implies that $M_P(X)=\pi(b^2-a^2)$, for all P, and the proof of the theorem is complete.

The sets of the form X in the theorem are not necessarily spherical, hence

the title of this note is justified. In particular, $X=C$ is a nonspherical topological 3-cell (and therefore connected, with connected boundary, and contractible) with constant HA-measurements. Clearly, the restriction $C\subseteq X\subseteq B_{ab}$ on X leaves much freedom for choosing X.

Next, let b be close to a, and let $\epsilon>0$ be very small with respect to a; define D by

$$D = B_{ab} - \{(x, y, z) \mid y^2 + z^2 < \epsilon^2\}.$$

Then D is obtained from B_{ab} by removing two symmetrically located cylindrical plugs of diameter 2ϵ. For every plane P there exists a parallel plane P' which meets the inner sphere of B_{ab} and misses both the plugs; $A(P'\cap D)=\pi(b^2-a^2)$, by our lemma, and since $D\subset B_{ab}$, $M_P(D)=\pi(b^2-a^2)$. As a result, D has constant HA-measurements. D is a centrally symmetric and nonspherical body which is topologically a solid torus, hence noncontractible and connected with a connected boundary.

A similar body D' may be obtained from B_{ab} by removing more than one pair of symmetrically located plugs; one can choose D' to be centrally symmetric such that $D\subset D'\subset B_{ab}$, or else take the plugs to be far apart from each other.

In connection with the use of HA-measurements in solid-state physics (see [1]), V. Klee asked (in private communication) whether there exists a body B with constant HA-measurements which is nonspherical and yet is centrally symmetric, has connected boundary, and contains a small ball centered at its center of symmetry.

Our affirmative answer to this question is as follows: let a, b, c, and t be such that $a\gg b-a\geqq 1$, $t\ll 1$, and $b^2-a^2=c^2-b^2$. The cone given by $y^2+z^2=t^2x^2$ divides E^3 into two parts:

$$K_1 = \{(x, y, z) \mid y^2 + z^2 \leqq t^2x^2\} \quad \text{and} \quad K_2 = \{(x, y, z) \mid y^2 + z^2 \geqq t^2x^2\}.$$

The example F we have in mind is given by

$$F = B_{0a} \cup (B_{ab} \cap K_1) \cup (B_{bc} \cap K_2).$$

The body F is symmetric with respect to the origin, has a boundary which is connected (though not a 2-manifold), and contains the ball B_{0a} of radius a centered at the origin.

It is not at all obvious (although it is true) that F has constant HA-measurements. It is easy to show that for all planes P, passing through the origin,

$$A(P \cap F) = \pi(a^2) + \lambda\pi(b^2 - a^2) + (1 - \lambda)\pi(c^2 - b^2) = \pi b^2$$

for an appropriate λ with $0\leqq\lambda<1$. Hence $A(P\cap F)$ is independent of P for planes P passing through the origin. It is much more complicated to show that $A(P\cap F)<\pi b^2$ for every plane P that does not pass through the origin. We have carried out the computations in the particular case where $a=9$, $b=10$, $c=\sqrt{119}$, and $t=\sqrt{11}/33$. The details are omitted, but will be supplied upon request.

This example originated from the following observations:

1. If P is close to the yz-plane and does not pass through the origin, then $A(P\cap F)<\pi b^2$ follows by using our lemma.

2. If P is close to the x-axis, and intersects the cone $y^2+z^2=t^2x^2$ in a hyperbola, then the parts $P\cap B_{ab}\cap F$ and $P\cap B_{bc}\cap F$ of $P\cap F$ are partly bounded by the hyperbola, which implies that $A[P\cap F\cap(B_{ab}\cup B_{bc})]<\pi(c^2-a^2)$; as a result, $A(P\cap F)<\pi(c^2-b^2)+\pi a^2=\pi b^2$.

Reference

1. V. Klee, Is a body spherical if its HA-measurements are constant? this MONTHLY, 76 (1969) 539–542.

A NOTE ON END-SETS

A. M. BRUCKNER AND J. G. CEDER, University of California, Santa Barbara

1. Introduction. A subset of Euclidean space is called an *end-set* if it is equal to the set of endpoints of a family of pairwise disjoint closed line segments. In [2] V. Klee and M. Martin posed the problem of whether there exists a compact end-set of positive Lebesgue measure in R^n where $n\geqq 3$. That the answer is negative when $n=2$ follows from an argument given in [2]. In this note we answer this question affirmatively whenever $n\geqq 4$. The question remains open for $n=3$.

2. The Case $n\geqq 4$.

THEOREM. *If $n\geqq 4$ there exists a compact end-set in R^n of positive measure.*

Proof. It suffices to prove the theorem for $n=4$. In [4] Nikodym constructed a compact set N in R^2 of positive Lebesgue measure with the property that corresponding to each $x\in N$ there is a line $L(x)$ passing through x such that $N\cap L(x)=\{x\}$. Consider the set $N\times N=N^2$ in R^4. It follows that through each $z\in N^2$ there passes a (two-dimensional) plane $P(z)$ such that $P(z)\cap N^2=\{z\}$. We let S denote the hemispherical shell $\{x:\|x\|=1,\ x_1\geqq 0\}$. Without loss of generality we may assume that N^2 is a subset of the interior of the convex hull of S.

First we shall construct a family of disjoint rays emanating from the points of N^2. Let N^2 be well-ordered by the ordinal c (i.e., the first ordinal equivalent to R^1) so that $N^2=\{z_\alpha:\alpha<c\}$. Let $T(z_0)$ be any ray in the plane $P(z_0)$ which emanates from z_0 and intersects S. Now assume we have chosen for each $\alpha<\beta$ a ray $T(z_\alpha)$ emanating from z_α and intersecting S such that $\gamma<\alpha<\beta$ implies $T(z_\gamma)\cap T(z_\alpha)=\varnothing$. Since

$$\operatorname{card}(P(z_\beta)\cap \bigcup\{T(z_\alpha):\ \alpha<\beta\})\leqq \operatorname{card}\beta<c,$$

there exists a line L in the plane $P(z_\beta)$ which intersects S and which misses the set $\bigcup\{T(z_\alpha):\alpha<\beta\}$. At least one of the two rays lying in L and emanating from z_β intersects S. Let $T(z_\beta)$ be such a ray. Clearly the collection of rays

$$\{T(z_\alpha): \alpha < c\} = \{T(z): z \in N^2\}$$

is disjoint, and the set N^2 comprises "half" of the set of end points of this collection of rays. If we adjoin to N^2 the points on the surface of the hemisphere S at which the collection of rays $\{T(z): z \in N^2\}$ intersect, then we obtain an end-set of positive measure which, however, need not be compact.

In order to enlarge N^2 to a compact end-set we proceed as follows: Let f be the reflection about the plane $x_1 = 2$; that is, $f(z) = (4 - z_1, z_2, z_3, z_4)$. We claim that the set X consisting of $N^2 \cup S$ together with its reflected image under f is a compact end-set of positive measure.

Clearly X is compact and of positive measure. To complete the proof it suffices to construct a family of disjoint line segments $\{V(x): x \in X\}$ whose ends comprise X. For $x \in N^2$ let $V(x)$ be the line segment contained in the ray $T(x)$ which joins x to the point, denoted by x', in $S \cap T(x)$. For $x \in f(N^2)$ let $V(x)$ be the line segment joining x to the point $f(w')$, where $f(w) = x$. For $x \in S - \{w': w \in N^2\}$ let $V(x)$ be the line segment joining x to $f(x)$. Finally, for

$$x \in f(S) - \{f(w'): w \in N^2\} \qquad \text{let } V(x) = V(f^{-1}(x)).$$

It can be easily checked that the family $\{V(x): x \in X\}$ is disjoint and has X as its set of end points.

3. Further remarks. The problem is unsolved for $n = 3$ and seems difficult. However, there does exist a compact end-set of positive Lebesgue outer measure in R^3, and there exists a compact set of positive Lebesgue measure in R^3 which is comprised of the end points of a family of pairwise disjoint arcs. For the former example, if f is a function from R^2 into R whose graph has outer measure positive, then the graph of f union the graph of the function $f+1$ yields the desired end-set. For the latter example, let C be a nowhere dense perfect subset of R^1 having positive measure. It is known [3, p. 173] that there exists a simple closed curve J which contains $C \times C$. Therefore J has positive measure in R^2. There exists a homeomorphism of the plane which carries J onto the unit circle. Since the unit circle is the set of ends of a family of disjoint arcs the same is true of J. This example, due to Klee and Martin, may be extended in the obvious way to obtain a homeomorphism of R^3 into R^3 such that the image of the unit sphere contains $C \times C \times C$ which has positive measure in R^3. The result follows as before.

A likely candidate for a compact end-set of positive measure in R^3 is the set $N \times C$, where N is the Nikodym set of the theorem and C is a suitably chosen nowhere dense perfect subset of R^1 having positive measure. One promising method of attack is the following: well-order N as $\{z_\alpha: \alpha < c\}$ and by induction define on $\{z_\alpha\} \times C$ a "comb" of disjoint rays emanating from the points of $\{z_\alpha\} \times C$ in such a way that these combs never intersect one another. Then, employing the ideas in the proof of the theorem, one could construct a bigger compact set which is an end-set of positive measure. Unfortunately we are unable to construct such a family of "combs". Some results about combs which appear to be related to the problem can be found in [1].

One could also obtain a compact end-set of positive measure in R^3 using the ideas in the proof of the theorem, provided there exists a compact set A of positive measure in R^3 with the following property: associated with each $x \in A$ there exists a family of lines $\mathfrak{L}_x$ passing through x such that $L \cap A = \{x\}$ whenever $L \in \mathfrak{L}_x$ and $\bigcup \mathfrak{L}_x$ is not contained in the union of less than c planes. This would be the case, for example, if the Nikodym set N could be replaced by a set M of positive measure with the property that for each $x \in M$ there exists c distinct lines $\{L_\alpha(x) : \alpha < c\}$ such that $L_\alpha(x) \cap M = \{x\}$ for each x and α. Then $M \times C$ would have the above property where C is any nowhere dense perfect subset of R^1.

The first author was supported in part by NSF grant GP-8253.

References

1. A. Bruckner and J. Ceder, Combs, to appear in Rev. Math. Pures et Appl.
2. V. Klee and M. Martin, Semi-continuity of the face-function of a convex set, Comment. Math. Helv. to appear.
3. K. Kuratowski, Topology, vol. II, Academic Press, New York, 1968.
4. O. Nikodym, Sur la mesure des ensembles plans dont tous les points sont rectilinéairement accessibles, Fund. Math., 10 (1927) 116–168.

A CLASS OF INTEGRATION FORMULAS

R. W. HAMMING, Bell Telephone Laboratories, Murray Hill

We can integrate $\int_0^1 f(x)dx = \int_0^1 y'(x)dx = y_1 - y_0$ using any of the following formulas:

$$y_1 = y_0 + \frac{y'(\theta_0)}{1!},$$

$$y_1 = y_0 + \frac{1}{2}(y_1' + y_0') - \frac{y'''(\theta_1)}{12},$$

$$y_1 = y_0 + \frac{1}{2}(y_1' + y_0') - \frac{1}{12}(y_1'' - y_0'') + \frac{y^{(5)}(\theta_2)}{720},$$

where each θ lies on the interval $0 < \theta < 1$. The existence of such θ's is guaranteed by the mean value theorem.

The form of the general formula is

$$y_1 = y_0 + \sum_{k=1}^{N} A_k^N \left(y_1^{(k)} - (-1)^k y_0^{(k)}\right) + \frac{E_{2N+1} y^{(2N+1)}(\theta)}{(2N+1)!} \tag{1}$$

and is *not* the truncated Euler-Maclaurin expansion. (For related material see: P. M. Hummel and C. L. Seebeck, this MONTHLY, 56 (1949) 243–247, *A generalization of Taylor's expansion*, and 58 (1951) 383–89, *A new interpolation formula*.)

These formulas may be used for integration as well as for correctors in pre-

dictor-corrector methods. Their potential value in the numerical integration of ordinary differential equations comes from the observation that often the majority of the machine time is used in the computation of radicals, logs, exponential, and trigonometric functions, and that the higher derivatives introduce no essentially new ones. Thus the higher derivatives are obtained relatively cheaply.

The arrangement of the material follows the main lines of its discovery rather than that of minimal, elegant presentation. This is done in the belief that the methods of mathematics are more important than the specific results, especially in these days of almost infinite knowledge. It is hoped that the spirit of the paper, first ask what is wanted and why, then find the first few specific cases, and finally explore the general case (with its curious bypaths through generating functions and contour integration), accurately reflects how mathematics is often done in practice.

One method for deriving the A_k^N is to first make the formula (1) (without the error term) exact for $y=1, x, x^2, \cdots, x^{2N}$. To preserve symmetry it is better to use the equivalent functions

$$y = 1, \quad (x - \tfrac{1}{2}), \quad (x - \tfrac{1}{2})^2, \cdots, (x - \tfrac{1}{2})^{2N}.$$

The first few cases are easily done, and are given in Table 1 along with the corresponding influence functions $G_N(s)$. Here $A_0^N = 1$ is the coefficient of y_0. (See for example: R. W. Hamming, *Numerical Methods for Scientists and Engineers*, McGraw-Hill, 1962, p. 144; F. B. Hildebrand, *Introduction to Numerical Analysis*, McGraw-Hill, 1956, p. 164; W. E. Milne, *Numerical Calculus*, Princeton U. Press, 1949, p. 111, or A. Ralston, *A First Course in Numerical Analysis*, McGraw-Hill, 1965, p. 167.)

TABLE 1.

$N \backslash k$	A_0^N	A_1^N	A_2^N	A_3^N	A_4^N	E_{2N+1}	Error Term Coefficient	$(2N)!\, G_N(s)$
0	1					1	1	1
1	1	1/2				−1/2	−1/12	$[-s(1-s)]$
2	1	1/2	−1/12			1/6	1/720	$[-s(1-s)]^2$
3	1	1/2	−1/10	1/120		−1/20	$\frac{-1}{100{,}800}$	$[-s(1-s)]^3$
4	1	1/2	−3/28	1/84	−1/1680	1/70	$\frac{1}{25{,}401{,}600}$	$[-s(1-s)]^4$

The general case, of theoretical interest mainly, can be done in a closed form as follows. Using

$$y = (x - \tfrac{1}{2})^m \qquad (0 \leqq m \leqq 2N)$$

in the formula, and adopting the factorial notation $x^{(n)} = x(x-1) \cdots (x-n+1)$, we get

$$(\tfrac{1}{2})^m = (-\tfrac{1}{2})^m + \sum_{k=1}^{N} A_k^N m^{(k)} [(\tfrac{1}{2})^{m-k} - (-1)^k (-\tfrac{1}{2})^{m-k}].$$

For m an even number, the formulas are satisfied for any A_k^N; we need consider only $m = 1, 3, 5, \cdots, 2N-1$. For these values of m we can rearrange the expression in the form

$$\sum_{k=1}^{N} A_k^N m^{(k)} 2^k = 1 \qquad (m = 1, 3, \cdots, 2N - 1).$$

It is not immediately obvious how to solve these equations (in the general case); instead we observe that the regularity of $G_N(s)$ suggests an approach—we calculate $G_N(s)$ two ways. One way we *guess* that

$$(2N)!\, G_N(s) = [-s(1 - s)]^N.$$

The other way we base on using $(x-s)_+^{2N}$ in the formula, and obtain

$$\begin{aligned}(2N)!\, G_N(s) &= (1 - s)^{2N} - \sum_{k=1}^{N} A_k^N (2N)^{(k)} (1 - s)^{2N-k} \\ &= (1 - s)^N \left\{ (1 - s)^N - \sum_{k=1}^{N} A_k^N (2N)^{(k)} (1 - s)^{N-k} \right\}.\end{aligned}$$

We therefore equate the curly bracket to $(-s)^N$:

$$(-s)^N \equiv (1 - s)^N - \sum_{k=1}^{N} A_k^N (2N)^{(k)} (1 - s)^{N-k}.$$

Set $1 - s = t$:

$$(t - 1)^N \equiv \sum_{k=0}^{N} (-1)^k C(N, k) t^{N-k} \equiv t^N - \sum_{k=1}^{N} A_k^N (2N)^{(k)} t^{N-k}.$$

Equating like powers in t we get $(-1)^k C(N, k) = -A_k^N (2N)^{(k)}$, or

$$A_k^N = (-1)^{k-1} \frac{C(N, k)}{(2N)^{(k)}} \qquad (k = 1, 2, \cdots, N).$$

To check our guess we put these into their defining equations and prove that the equations are now identities. We have

$$\sum_{k=1}^{N} (-1)^{k-1} \frac{C(N, k)}{(2N)^{(k)}} m^{(k)} 2^k \equiv 1,$$

or

$$\sum_{k=0}^{N} (-2)^k \frac{C(N, k) m^{(k)}}{(2N)^{(k)}} \equiv 0, \qquad (m = 1, 3, \cdots, 2N-1).$$

This can be rewritten as

$$\sum_{k=0}^{N} (-2)^k \frac{N!}{k!(N-k)!} \frac{m!}{(m-k)!} \frac{(2N-k)!}{(2N)!} \equiv 0,$$

$$\frac{N!N!}{(2N)!} \sum_{k=0}^{N} (-2)^k \frac{(2N-k)!}{N!(N-k)!} \frac{m!}{k!(m-k)!} = 0,$$

$$\sum_{k=0}^{N} (-2)^k C(2N-k, N) C(m, k) = 0 \qquad (m = 1, 3, \cdots, 2N-1).$$

Set $k = N - j$:

$$\sum_{j=0}^{N} (-2)^{N-j} C(N+j, N) C(m, N-j) = 0.$$

But $C(N+j, N) = C(N+j, j) = C(-(N+1), j)(-1)^j$, so we have

$$\sum_{j=0}^{N} (2)^{N-j} C(-(N+1), j) C(m, N-j) = 0.$$

This suggests the generating function

$$\frac{(1+2x)^m}{(1+x)^{N+1}} = \sum_{n,j} C(-(N+1), j) C(m, n-j) 2^{n-j} x^n.$$

We need to pick out the coefficient of x^N and show that it is zero for $m = 1, 3, \cdots, 2N-1$.

By contour integration about $z = 0$, the coefficient of z^N is given by

$$\frac{1}{2\pi i} \oint \frac{(1+2z)^m}{(1+z)^{N+1} z^{N+1}} dz.$$

Put $z = t - \frac{1}{2}$; the integral becomes

$$\frac{1}{2\pi i} \oint \frac{(2t)^m}{(t^2 - \frac{1}{4})^{N+1}} dt \qquad (\text{around } t = \tfrac{1}{2}).$$

Deform the contour to the iy axis and toward infinity. Since $m \leqq 2N-1 < 2(N+1) - 2$, the contribution along the semicircle approaches zero. Set $t = iy$:

$$(\text{const.}) \int_{-\infty}^{\infty} \frac{y^m}{(y^2 + \frac{1}{4})^{N+1}}\, dy = 0,$$

since m is odd and less than $2N$; the identities have been proved.

To find the error term we use our (now verified) guess

$$(2N)!\, G_N(s) = [-s(1 - s)]^N \qquad (0 \leqq s \leqq 1).$$

The error term is, because $G_N(s)$ is of constant sign,

$$\frac{1}{(2N)!} \int_0^1 y^{(2N+1)}(s) G_N(s) ds = \frac{1}{(2N)!} y^{(2N+1)}(\theta) \int_0^1 G_N(s) ds.$$

Now $\int_0^1 [s(1-s)]^N ds$ is easily integrated by the substitution $s = \sin^2\theta/2$:

$$\begin{aligned} \int_0^{\pi} \sin^{2N+1} \theta/2 \cos^{2N+1} \theta/2 \, d\theta &= \frac{1}{2^{N+1}} \int_0^{\pi} \sin^{2N+1} \theta \, d\theta \\ &= \frac{1}{2^N} \int_0^{\pi/2} \sin^{2N+1} \theta \, d\theta, \end{aligned}$$

the well-known Wallis integral. The error term, therefore, is

$$(-1)^N \left[\frac{N!}{(2N)!}\right]^2 \frac{1}{2N + 1} y^{(2N+1)}(\theta).$$

The table is thus easily extended if desired.

Thanks to E. N. Gilbert, R. Pinkham, and S. O. Rice.

ON THE GROUP OF UNITS OF A RING

S. Z. Ditor, Louisiana State University

There are several results in the literature which relate the structure of a ring (with identity) to that of its group of units. Gilmer [3] determines all finite commutative rings whose group of units is cyclic. Eldridge and Fischer [2] extend these results to artinian rings, and, in [1], Eldridge shows to what extent the structure of an artinian ring is determined by knowing that it has either a solvable, simple, nilpotent, supersolvable, torsion, or finitely generated group of units.

In a different vein, one can pose the following questions:

(1) Which groups can be the group of units of a ring?

(2) Which whole numbers can be the number of units of a ring?

The following theorem answers these questions for finite groups of *odd* order. Although the result can be obtained by using Theorem 1 of [1] (with the aid of the Feit-Thompson theorem concerning solvable groups), the direct proof offers

such a nice exercise in the use of classical structure theory that it deserves to be given.

THEOREM. (i) *A finite group G of odd order is the group of units of some ring if and only if G is abelian and is the finite direct product of cyclic groups G_i, where the order of each G_i is of the form $2^{k_i}-1$.*

(ii) *If G is the group of units of a ring R and if G is finite of odd order, then the subring $[G]$ of R generated by G is a finite direct sum of Galois fields of characteristic 2, namely*

$$[G] = \oplus_{i=1}^{r} GF(2^{k_i}).$$

Proof. We prove (ii), (i) being an immediate consequence of (ii) and the fact that the multiplicative group of a finite field is cyclic [4, V. I, p. 117].

Since G has odd order, $-1=1$; otherwise $\{-1, 1\}$ would be a subgroup of G of order 2. Hence the subring $[G]$ generated by G is a finite-dimensional algebra over $GF(2)$. Now, $[G]$ is a representation module of G over $GF(2)$ and since 2, the characteristic of $GF(2)$, does not divide the order of G, Maschke's theorem [4, V.II, p. 179] implies that $[G]$ is semisimple. By the Wedderburn-Artin theorem [4, V.II, p. 156], $[G]$ is the finite direct sum of rings A_i, where each A_i is the full ring of $n_i \times n_i$ matrices (for some n_i) over a division ring D_i, $i=1, \cdots, r$. By Wedderburn's theorem [4, V.II, p. 203], the finite division rings D_i must in fact be fields, and since $-1=1$ in $[G]$, each D_i is a Galois field of characteristic 2. Each $n_i=1$ because, as is readily verified, the ring of $n\times n$ matrices over a finite field of s elements has precisely $(s^n-1)(s^n-s)\cdots(s^n-s^{n-1})$ units (matrices whose rows are linearly independent), and, when s is even, this number is odd if and only if $n=1$.

COROLLARY. *A prime power p^m is the number of units of some ring if and only if p is 2 or a Mersenne prime 2^q-1 for some number q.*

Proof. If $p=2$ (resp. 2^q-1), then $p+1=3$ (resp. 2^q), and the m-fold direct sum of $GF(3)$ (resp. $GF(2^q)$) has precisely p^m units. Conversely, if a group of prime power order p^m is the group of units of a ring and if $p\neq 2$, then, by the above theorem, p^m is a product of numbers of the form 2^k-1. Therefore, for some positive integers n and k, $p^n=2^k-1$. For n even, $p^n-1=(p-1)(p^{n-1}+\cdots+p+1)$ is divisible by 4, whereas for $k\neq 1$, 2^k-2 is not. Hence n is odd and $p+1$ divides $p^n+1=2^k$, so that $p+1$ is a power of 2.

References

1. K. E. Eldridge, On ring structures determined by groups, Proc. Amer. Math. Soc., 23 (1969) 472–477.

2. K. E. Eldridge and I. Fischer, D. C. C. rings with a cyclic group of units, Duke Math. J., 34 (1967) 243–248.

3. R. Gilmer, Finite rings having a cyclic multiplicative group of units, Amer. J. Math., 85 (1963) 447–452.

4. B. L. van der Waerden, Modern Algebra, Volumes I and II, Ungar, New York, 1949 and 1950.

A NOTE ON ORDER STATISTICS

A. G. Konheim, T. J. Watson Res. Center, IBM

Let $X_1, X_2, \cdots$ be independent, identically distributed random variables with distribution function F. Let $X_{1,n} \leqq X_{2,n} \leqq \cdots \leqq X_{n,n}$ be the order statistics of $X_1, X_2, \cdots, X_n$. We assume that $E(|X_1|) < \infty$, which implies that $E(|X_{i,n}|) < \infty$, $1 \leqq i \leqq n$. It is known that the triangular array of numbers

$$\{E(X_{i,n}):\ 1 \leqq i \leqq n,\ 1 \leqq n < \infty\}$$

determines F. In particular (cf. [2]) the distribution function F_n which assigns equal weights to the n numbers $\{E(X_{i,n}): 1 \leqq i \leqq n\}$ converges to F. The purpose of this note is to prove the following stronger result:

THEOREM. *The numbers* $\{E(X_{n,n}): 1 \leqq n < \infty\}$ *determine* F.

Proof: Let F and G be distribution functions,

$$\int_{-\infty}^{\infty} |x|\, F(dx) < \infty, \qquad \int_{-\infty}^{\infty} |x|\, G(dx) < \infty,$$

and suppose $\mu_n(F) = \mu_n(G)\,(1 \leqq n < \infty)$ with μ_n defined by $\mu_n(F) = \int_{-\infty}^{\infty} x\,dF^n(x)$. Since $E(|X_{n,n}|) < \infty$, integration by parts yields

$$\mu_n(F) = -\int_{-\infty}^{0} F^n(x)dx + \int_{0}^{\infty} (1 - F^n(x))dx.$$

Let λ be Lebesgue measure and $\nu_F = \lambda F^{-1}$. We then have

$$\mu_{n+1}(F) - \mu_n(F) = \int_0^1 y^n(1 - y)\nu_F(dy),$$

so if we set $\omega_F(dy) = y(1-y)\nu_F(dy)$, we may conclude that

$$\int_0^1 y^k \omega_F(dy) = \int_0^1 y^k \omega_G(dy) \qquad (0 \leqq k < \infty).$$

But ω_F and ω_G are supported on a finite interval, and consequently [**1**; pp. 222–224] $\omega_F = \omega_G$. This implies $F = G$.

I would like to acknowledge with thanks several conversations with Dr. Alan J. Hoffman on this problem.

This research was partially supported by the United States Air Force under Contract No. AF49(638)-1682.

References

1. William Feller, An Introduction to Probability Theory and its Applications, Volume II, Wiley, New York, 1970.

2. Wassily Hoeffding, On the distribution of the expected values of the order statistics, Ann. Math. Stat., 24 (1953) 93–100.

HILBERT SPACE PROBLEM FOUR

G. G. JOHNSON, Virginia Polytechnic Institute

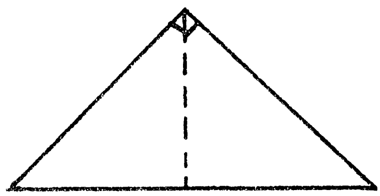
Problem four in [1] reads as follows: Construct, for every infinite dimensional Hilbert space, a simple continuous curve with the property that every two nonoverlapping chords of it are orthogonal.

The following is an elementary solution in the sense that only the abstract definition of Hilbert space is used.

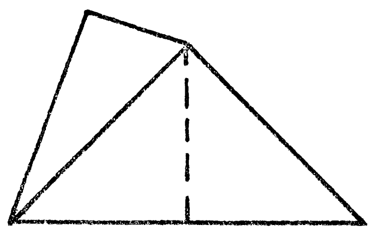
To have a picture in mind as to how this is to be done consider the following sequence of diagrams:

Start with an interval of length one.

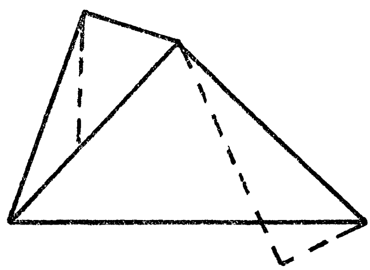
Construct an isosceles triangle with this interval as base and a right angle at the vertex.



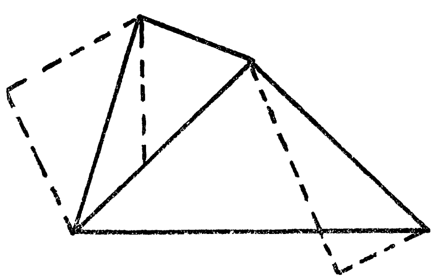
Fold out the left triangle such that it is orthogonal to the plane of the original.



Fold out the right triangle such that it is orthogonal to the space generated by the left triangle and the original triangle.



Continue this process with each of the two triangles folded out, taking care that each triangle folded out is orthogonal to all preceding triangles.



Continue in this manner.

Note that the vertices remain fixed and hence we have a countable vertex collection V each point of which is a limit point of the collection. The closure of V will be the desired arc.

Having such a picture in mind let us proceed to an analytic description.

Let $\{\phi_i\}_{i=1}^{\infty}$ be an orthonormal sequence and define for each positive integer n, a function f_n as follows:

$$f_1(x) = x\phi_1 \qquad \text{if } x \in [0, 1],$$

$$f_{n+1}(x) = 2^{n-1}\left[x - \frac{j}{2^{n-1}}\right]\left[f_n\left(\frac{j+1}{2^{n-1}}\right) - f_n\left(\frac{j}{2^{n-1}}\right) + (\sqrt{2})^{1-n}\phi_{2^{n-1}+j+1}\right]$$
$$+ f_n\left(\frac{j}{2^{n-1}}\right) \qquad \text{if } x \in [0, 1] \cap \left(\frac{2j}{2^n}, \frac{2j+1}{2^n}\right)$$

and

$$f_{n+1}(x) = 2^{n-1}\left[x - \frac{2j+1}{2^n}\right]\left[f_n\left(\frac{j+1}{2^{n-1}}\right) - f_n\left(\frac{j}{2^{n-1}}\right) - (\sqrt{2})^{1-n}\phi_{2^{n-1}+j+1}\right]$$
$$+ f_{n+1}\left(\frac{2j+1}{2^n}\right) \qquad \text{if } x \in [0, 1] \cap \left(\frac{2j+1}{2^n}, \frac{2j+2}{2^n}\right).$$

For notational purposes let

$$\theta(n, j) = (\sqrt{2})^{1-n}\phi_{2^{n-1}+j+1} \quad \text{and} \quad D(n, j) = f_n((j+1)/2^{n-1}) - f_n(j/2^{n-1})$$

for $n = 1, 2, \cdots$ and $0 \leqq j < 2^{n-1}$.

By simple induction arguments we have for each positive integer n that:

1. $(f_n(x), \phi_1) = x \qquad \text{for all } x \text{ in } [0, 1],$
2. $\|D(n, j)\| = (\sqrt{2})^{1-n} \qquad \text{if } 0 \leqq j < 2^{n-1},$
3. $\|\theta(n, j)\| = (\sqrt{2})^{1-n} \qquad \text{if } 0 \leqq j < 2^{n-1}.$

The sequence $\{f_n\}_{n=1}^{\infty}$ has the following four properties:

(1) $f_{n+1}(t/2^{n-1}) = f_n(t/2^{n-1}),$

(2) $\|f_{n+1}(x) - f_n(x)\| \leqq (\sqrt{2})^{1-n},$

(3) $D(n, t) \perp D(n, s) \qquad \text{if } s \neq t,$

(4) $\|f_n(x) - f_n(y)\| \geqq |x - y|.$

From properties 2 and 4 it follows that $\lim_{n\to\infty} f_n(x) = f(x)$ exists on $[0, 1]$ and that f is a homeomorphism.

Suppose $0 \leqq a < b \leqq c < d \leqq 1$ and each of a, b, c, and d is a dyadic rational; then by properties 1, 3, and the above comments, it follows that $(f(b) - f(a))$ is orthogonal to $(f(d) - f(c))$. Since the inner product is continuous and the dyadic rationals are dense in $[0, 1]$ it follows that if $0 \leqq a < b \leqq c < d \leqq 1$, then $(f(b) - f(a))$

is orthogonal to $(f(d)-f(c))$ and hence the arc $f([0,1])$ has the desired properties.

It now remains to show that the sequence $\{f_n\}_{n=1}^{\infty}$ has properties 1 through 4.

Property (1) is obvious from the construction. To establish property (2), consider first the case when $x\in[2j/2^n, (2j+1)/2^n]$. A straightforward calculation yields

$$\|f_{n+1}(x)-f_n(x)\| = \|2^{n-1}(x-j)\theta(n,j)\| = (\sqrt{2})^{1-n}\,|\,2^{n-1}(x-j)\,|\ \leqq (\sqrt{2})^{1-n}.$$

The case when $x\in((2j+1)/2^n, (2j+2)/2^n)$ is settled in a similar fashion.

To establish property (3), note that the range of f_n is contained in the subspace spanned by $\{\phi_i : 1\leqq i\leqq 2^{n-1}\}$. Also note that

$$D(n+1, 2k) = [D(n,k) + (\sqrt{2})^{1-n}\phi_{2^{n-1}+k+1}]/2$$

and

$$D(n+1, 2k+1) = [D(n,k) - (\sqrt{2})^{1-n}\phi_{2^{n-1}+k+1}]/2.$$

By definition of f_2 it is clear that $D(2,s)$ is orthogonal to $D(2,t)$ if $s\neq t$. Suppose now that n is a positive integer such that if $1<m\leqq n$, then $D(m,s)$ is orthogonal to $D(m,t)$ if $s\neq t$. Consider now $(D(n+1,2k), D(n+1,2l))$, $(D(n+1,2k), D(n+1,2l+1))$ and $(D(n+1,2k), D(n+1,2k+1))$ where $k\neq l$. By the above remarks it readily follows that each of the first two inner products is zero. In the third inner product we have

$$\begin{aligned}\tfrac{1}{4}[(D(n,k), D(n,k)) &+ (D(n,k), (\sqrt{2})^{1-n}\phi_{2^{n-1}+k+1}) - ((\sqrt{2})^{1-n}\phi_{2^{n-1}+k+1}, D(n,k)) \\ &- ((\sqrt{2})^{1-n}\phi_{2^{n-1}+k+1}, (\sqrt{2})^{1-n}\phi_{2^{n-1}+k+1})] = \tfrac{1}{4}[\|D(n,k)\|^2 - 2^{1-n}] \\ &= \tfrac{1}{4}[2^{1-n}-2^{1-n}] = 0.\end{aligned}$$

Thus the inductive step is completed and property (3) is established.

Property (4) is established by the following simple computation:

$$\|f_n(x)-f_n(y)\| = \|f_n(x)-f_n(y)\|\cdot\|\phi_1\| \geqq |\,(f_n(x)-f_n(y), \phi_1)\,| = |\,x-y\,|.$$

This then completes the verification of the four properties and hence the example.

It perhaps is of some interest to note that the sequence $\{f_n\}_{n=1}^{\infty}$ has the following formulation: $f_1(x)=x\phi_1$ for $x\in[0,1]$ and for each positive integer n

$$f_{n+1}(x) = f_n(x) + 2^{n-1}[x-2j/2^n]\theta(n,j)$$

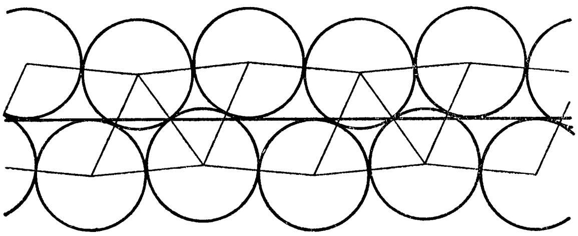
if $x\in[2j/2^n, (2j+1)/2^n]$, and $0\leqq j\leqq 2^{n-1}$, and

$$f_{n+1}(x) = f_n(x) + 2^{n-1}[(2j+2)/2^n - x]\theta(n,j)$$

if $x\in((2j+1)/2^n, (2j+2)/2^n)$ and $0<j<2^{n-1}$.

References

1. Paul R. Halmos, A Hilbert Space Problem Book, Van Nostrand, Princeton, N. J., 1967.
2. G. G. Johnson, A crinkled arc, Proc. Amer. Math. Soc., 25 (1970) 375–376.



References

1. Problem 1 of the M. Schweitzer mathematical competition 1966. Mat. Lapok, 18 (1967) 108–109.

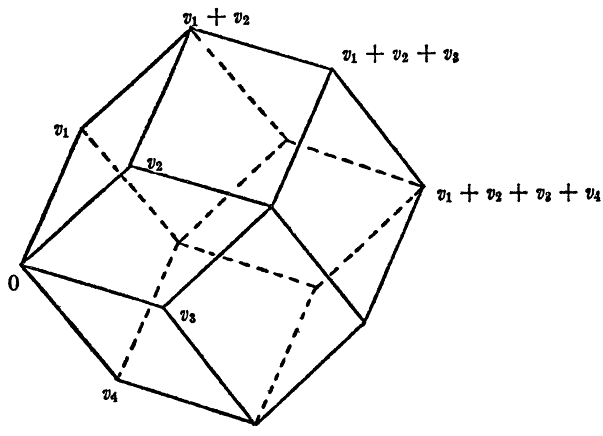
2. J. Horváth, Sphere-packings in 3- and more-dimensional Euclidean spaces (in Hungarian). (1970, manuscript.)

THE ZONOID PROBLEM

E. D. Bolker, Bryn Mawr College

Let $v_1, \cdots, v_N$ be vectors in Euclidean n-space R^n; think of them as arrows with their tails at the origin. Translate v_1 parallel to itself along v_2, forming the parallelogram with vertices 0, v_1, v_2, and v_1+v_2. Now translate that parallelogram along v_3, the resulting figure along v_4, and so on. Whenever v_i is in the linear span of its predecessors, the translation along v_i will cause the figure to pass through itself. The polytope Z we have just constructed is called a *zonotope*. Algebraically,

$$Z = \sum_{i=1}^{N} [0, v_i] = \left\{ \sum_{i=1}^{N} \lambda_i v_i \colon 0 \leqq \lambda_i \leqq 1 \right\},$$



a *sum of segments*. Z is the linear image in R^n of the unit cube

$$\{(\lambda_1, \cdots, \lambda_N): 0 \leqq \lambda_i \leqq 1\}$$

in R^N under the map which takes the ith basis vector in R^N to v_i in R^n. When $N=4$ and $n=3$, the v_i can be chosen so that Z is the classical rhombic dodecahedron, as shown in the figure.

The study of zonotopes began at least as early as 1611 when Kepler wrote, in [5, p. 11]: "These rhombi (observed in a honeycomb) put it into my head to embark on a problem of geometry: whether any body . . . could be constructed with nothing but rhombi (for faces)." The Russian crystallographer E. S. Federov studied 3-dimensional zonotopes in 1885 [4] in order to solve packing problems. In 1933, A. D. Alexandrov characterized zonotopes as those polytopes whose opposite faces are centrally symmetric [1].

A closed, centrally symmetric convex body (henceforth abbreviated as *body*) is called a *zonoid* if it can be approximated arbitrarily closely by zonotopes. The *zonoid problem* is to find a generalization of Alexandrov's theorem which gives intrinsic geometric criteria for deciding when a body is a zonoid. Blaschke, in 1916 [2, p. 145 in 2nd ed.] seems to have been the first to formulate the problem. An extensive bibliography and proofs of most of the assertions in this note can be found in [3].

Every 2-dimensional body is a zonoid, but in dimensions $n \geqq 3$ the zonoids are nowhere dense in the set of bodies, since every face of a zonoid is a zonoid and hence a body with at least one triangular face is not a zonoid. In particular, the octahedron is not a zonoid. However, criteria depending on facial structure cannot solve the zonoid problem since most bodies are faceless. The three dimensional zonoid problem seems to be just as hard as the n-dimensional one. The partial solutions known are, unfortunately, neither intrinsic nor geometric. For example, a body is a zonoid just when it is a weak *-continuous linear image of the unit ball of L^∞, which is a kind of infinite dimensional cube, and hence, according to results of Lindenstrauss and Liapounoff, when it is the range of a vector valued measure. But neither of these characterizations helps one decide whether a particular body one is interested in is a zonoid. Dually, a body K centered at the origin is a zonoid just when its *polar* $K^* = \{y \in R^n: \langle x, y\rangle \leqq 1$ for all $x \in K\}$ is a central section of the unit ball of L^1. For example, the octahedron is the polar of the cube and is the unit ball of the subspace of $L^1((0, 3))$ spanned by the characteristic functions of the intervals $(0, 1)$, $(1, 2)$ and $(2, 3)$. It is an interesting exercise to find four functions in L^1 which span a subspace whose unit ball is the polar of the rhombic dodecahedron.

The *support function* Ψ_K of a body K is given by $\Psi_K(y) = \sup\{\langle x, y\rangle : x \in K\}$. When $\|y\| = 1$, Ψ_K measures the distance from the origin to the support plane of K perpendicular to y and it determines K, because K is closed and convex. The body K is a zonoid if and only if there is a positive measure μ on the $n-1$-sphere S for which

$$\Psi_K(y) = \frac{1}{2}\int_S |\langle x, y\rangle|\, d\mu(x). \tag{1}$$

Then K is the "continuous sum" of segments corresponding to unit vectors, where the sum is weighted by the measure μ. The situation described in the opening paragraph results when μ is atomic and $\mu(\{v_i/\|v_i\|\}) = \|v_i\|$.

Let $x = (x_1, \cdots, x_n)$ and set $\Psi_p(x) = (\sum |x_i|^p)^{1/p}$. Then Ψ_p is the support function of the unit ball B_n^q of l_n^q, $\{x : \Psi_q(x) \leqq 1\}$, when $(1/p) + (1/q) = 1$. In [6, Section 63] P. Lévy proved a theorem, equivalent to the fact that B_n^q is a zonoid when $2 \leqq q \leqq \infty$. When $n \geqq 3$ and $1 \leqq q < 2$, B_n^q resembles the "octahedron" B_n^1 more than it does the "cube" B_n^∞; I conjecture that it is not a zonoid. H. Rosenthal (unpublished result) used techniques due to R. E. A. C. Paley [7] to prove the conjecture when $n = 3$ for $q \leqq 2 \log 7$ and for $q \leqq 2 \log n/\log 3n$ when $n > 3$. A stronger conjecture is the following characterization of the Euclidean norm: a body K and its polar are both zonoids just when K is an ellipsoid.

Finally, we present a sketch of a probabilistic construction of the zonoids which was discovered too late to be included in [3]. Perhaps it will help solve the zonoid problem.

Suppose $v_1, \cdots, v_N$ are vectors in space whose lengths sum to 1. Consider a particle initially at the origin which at any time will take a short step parallel to v_i with probability $\|v_i\|$. Then a possible, though not probable, path of length 1 will be piecewise straight and have length $\|v_i\|$ in direction v_i. The union of all such paths is precisely the zonotope $\sum [0, v_i]$ which therefore represents all the positions the particle might have occupied while following one of those paths. Unfortunately, those paths are not the results of a Markov process, for the particle must remember at each stage how much of its allotted length $\|v_i\|$ it has already used up in travelling parallel to v_i. Nevertheless there should be a sense in which the zonoid with support function given by (1) is the body swept out by most particles travelling a total distance of one unit, starting at the origin and at each instant choosing a new direction using the probability measure μ on the sphere.

References

1. A. D. Alexandrov, A theorem on convex polyhedra, Trudy Mat. Inst. Steklov., Sect. Math., 4 (1933) 87 (Russian).

2. W. Blaschke, Kreis und Kugel, Leipzig, 1916. 2nd ed. de Gruyter, Berlin, 1956.

3. E. D. Bolker, A class of convex bodies, Trans. Amer. Math. Soc., 145 (1969) 323–345.

4. E. S. Federov, Russisch-Kaiserliche Mineralogische Gesellschaft Verhandlungen (Zapiski Mineralogičeskih Obščestva) (2) 21 (1885) 1–279. Reprinted as Načala učeniya o figura (Elements of the study of figures), Izdat. Akad. Nauk SSSR, Moscow, 1953.

5. J. Kepler, The Six-cornered Snowflake, trans. Colin Hardie, Oxford, New York, 1966.

6. P. Lévy, Théorie de l'addition des fonctions aléatoires, Gauthier-Villars, Paris, 1937.

7. R. E. A. C. Paley, Some theorems on abstract spaces, Bull. Amer. Math. Soc., 42 (1936) 235–240.

CLASSROOM NOTES

EDITED BY DAVID DRASIN

Manuscripts for this Department should be sent to Robert Gilmer, Department of Mathematics, Florida State University, Tallahassee, FL 32306.

A CONTINUOUS LINEAR BIJECTION WITH DISCONTINUOUS INVERSE

R. K. WILLIAMS, Southern Methodist University

In most topology courses, one usually verifies that a one-to-one continuous function need not be a homeomorphism. Of course, if the domain of the function is compact and if the range is a subset of a Hausdorff space, then the function is a homeomorphism.

It is also known that a one-to-one continuous function which maps a connected open subset in E^n into E^n must be a homeomorphism, but that a one-to-one continuous function which maps a connected open subset of E^n into E^m, $m \neq n$, may not be a homeomorphism. (See [1], p. 156.)

The purpose of this paper is to give an example of a function f such that f is a one-to-one continuous linear function which maps a normed linear space *onto* itself, but such that f^{-1} is discontinuous at each point.

Let X be the set of all entire functions of a complex variable. Let addition and multiplication by complex scalars be defined in the usual way. Let a norm be defined on X by

$$\|f\| = \sup_{|z|=1} |f(z)|.$$

Standard arguments show that $\| \ \|$ is indeed a norm. For instance, if $\|f\| = 0$, then $f(z)$ is an entire function which is identically zero on $|z| = 1$. Since non-identically zero analytic functions have isolated zeros, $f(z) = 0$ for each z. Thus X is an infinite dimensional normed linear space over the complex field.

Define F on X by $F(f(z)) = f(z/2)$. Clearly F maps X onto itself in a one-to-one linear fashion. F is also uniformly continuous on X. For if $\epsilon > 0$,

$$f(z) = \sum_{n=0}^{\infty} a_n z^n, \qquad g(z) = \sum_{n=0}^{\infty} b_n z^n,$$

and $\|f - g\| < \epsilon$, then

$$\begin{aligned}\|F(f) - F(g)\| &= \sup_{|z|=1} \left| \sum_{n=0}^{\infty} (a_n - b_n)\left(\frac{z}{2}\right)^n \right| \leqq \sup_{|z|=1} \left| \sum_{n=0}^{\infty} (a_n - b_n) z^n \right| \\ &= \|f - g\| < \epsilon,\end{aligned}$$

where the first inequality is justified by the maximum modulus theorem.

We now show that F^{-1} is everywhere discontinuous. Clearly $F^{-1}(f(z)) = f(2z)$. Let $f_0 \in X$, and let $f_n(z) = f_0(z) + (z/2)^n$, $n = 1, 2, \cdots$. Then

$$\|f_n - f_0\| = \sup_{|z|=1} \left|\frac{z}{2}\right|^n = \frac{1}{2^n} \to 0,$$

but

$$\|F^{-1}(f_n) - F^{-1}(f_0)\| = \sup_{|z|=1} \left|\frac{2z}{2}\right|^n = 1.$$

Thus F^{-1} is discontinuous at each $f_0 \in X$, and hence F is not a homeomorphism.

It is easy to show that X is not complete. For example, the partial sums of $\sum_{n=0}^{\infty} (z/2)^n$ form a Cauchy sequence in X that does not converge in X. Hence X is not a Banach space. The Open Mapping Theorem implies that a one-to-one continuous linear transformation of a Banach space onto itself is a homeomorphism. Thus, the example above "works" because X is not complete.

Reference

1. Emil Artin and Hel Braun, Introduction to Algebraic Topology, (Translated by Erik Hemmingsen), Merrill Publ. Co., Columbus, Ohio, 1969.

LEXICOGRAPHIC PRODUCTS AND PERFECTLY NORMAL SPACES

J. E. VAUGHAN, University of North Carolina at Chapel Hill

1. Introduction. In this note we shall show that about half a dozen examples found in the literature of compact, perfectly normal, nonmetrizable spaces are all essentially the same (recall that a normal space is called perfectly normal if every closed set is a countable intersection of open sets). Our second purpose is to prove the theorem below which is a generalization of the fact (which is known; cf. [5] and [7]) that two of these spaces are actually homeomorphic. The two spaces in question are (i) the top and bottom line of the lexicographically ordered square: $[0, 1] \times_{\text{lex}} \{0, 1\}$, and (ii) the set of increasing functions defined on the closed unit interval $[0, 1]$ with values in the two-element discrete set $\{0, 1\}$ with the topology of pointwise convergence.

The notation and terminology used here are the same as in Kelley's book [6] with a few additions. Let X and Y be linearly (i.e., totally) ordered sets. Let X^{-1} denote the ordered set having the same elements as X, and order inverse to that of X. As in [3, p. 165] we write $X \times_{\text{lex}} Y$ for the lexicographically ordered product of X and Y (i.e., the set $X \times Y$ with the order defined by $(x, y) < (x', y')$ if either $x < x'$ or $x = x'$ and $y < y'$). We shall work with the order (i.e., interval) topology on ordered sets, and the topology of pointwise convergence (i.e., the relativized product topology) on subsets of Y^X, where Y^X denotes the set of all functions from X to Y. The ordered set X is called densely ordered if between every two distinct elements there exists a third.

THEOREM. *Let X be a densely ordered space and Y any ordered space having distinct first and last elements. Then*

(A) $X \times_{\text{lex}} Y$ *is homeomorphic to a subspace of the decreasing* (*i.e., order reversing*) *functions in* Y^X.

(B) $X^{-1} \times_{\text{lex}} Y$ *is homeomorphic to a subspace of the increasing* (*i.e., order preserving*) *functions in* Y^X.

2. Proof of the Theorem. We give a straightforward proof of part (A), and a similar proof can be given for part (B). Suppose y_0 and y_1 are respectively the first and last elements of Y. For each $(x, y) \in X \times Y$ we write

$$f_{x,y}(t) = \begin{cases} y_1 & \text{if } t < x \\ y & \text{if } t = x \\ y_0 & \text{if } t > x \end{cases}$$

and let $F = \{f_{x,y} : (x, y) \in X \times Y\}$. Recall that the product order $\leqq$ on Y^X is defined by $f \leqq g$ if and only if $f(x) \leqq g(x)$ for all $x \in X$.

LEMMA 1. $X \times_{\text{lex}} Y$ *is order isomorphic to* F *with the relative order induced by the product order.*

Proof. Define $\phi : X \times_{\text{lex}} Y \to F$ by $\phi(x, y) = f_{x,y}$. To see that ϕ is 1-1, we suppose $(x, y) < (z, w)$. Then if $x < z$, we have

$$\begin{cases} f_{x,y}(t) = y_1 = f_{z,w}(t) & \text{if } t < x \\ f_{x,y}(t) \leqq y_1 = f_{z,w}(t) & \text{if } t = x \\ f_{x,y}(t) = y_0 \leqq f_{z,w}(t) & \text{if } t > x \end{cases}$$

so $f_{x,y} \leqq f_{z,w}$. Since X is densely ordered, there exists $x < t < z$, and thus $f_{x,y}(t) = y_0 < y_1 = f_{z,w}(t)$ and $f_{x,y} < f_{z,w}$. If $x = z$ and $y < w$, then $f_{x,y}(x) < f_{z,w}(x)$, and in fact $f_{x,y} < f_{z,w}$. This shows not only that ϕ is 1-1 but also that ϕ is order preserving. Hence ϕ^{-1} is order preserving.

To complete the proof of part (A), it will suffice to show that the order topology on F induced by the product order is the same as the topology of pointwise convergence. This fact is established by the following two lemmas. We first recall that a subbase for the topology of pointwise convergence on F is $\{W(x, U) : x \in X, \text{ and } U \text{ open in } Y\}$, where $W(x, U) = \{f \in F : f(x) \in U\}$. Of course, we may consider only subbasic sets $U = (\leftarrow, y) = [y_0, y)$ or $U = (y, \rightarrow) = (y, y_1]$ for $y_0 < y < y_1$.

LEMMA 2. *The topology of pointwise convergence on* F *is a subset of the order topology on* F.

Proof. Since intervals of the form $(f_{x,y}, \rightarrow)$ and $(\leftarrow, f_{x,y})$ form a subbase for the order topology on F, it suffices to show for every $x \in X$ that

(a) $$W(x, (y, y_1]) = (f_{x,y}, \rightarrow) \qquad \text{for } y < y_1$$

(b) $$W(x, [y_0, y)) = (\leftarrow, f_{x,y}) \qquad \text{for } y > y_0.$$

We prove only (a) here. If $f_{z,w} \in W(x, (y, y_1])$, then $f_{z,w}(x) > y \geqq y_0$; so $z \geqq x$. If

$z>x$, then by Lemma 1 $f_{z,w}>f_{x,y}$. If $z=x$, then $f_{z,w}(x)=w>y$, and this implies $f_{z,w}>f_{x,y}$. In either case $f_{z,w}\in(f_{x,y}, \rightarrow)$. On the other hand, if $f_{z,w}\in(f_{x,y}, \rightarrow)$, then $(z, w)>(x, y)$ by Lemma 1. If $z>x$, then $f_{z,w}(x)=y_1$. If $z=x$ and $w>y$, then $f_{z,w}(x)=w>y$. In either case $f_{z,w}\in W(x, (y, y_1])$.

LEMMA 3. *The topology of pointwise convergence on* F *contains the order topology on* F.

Proof. It will suffice to show that each subbasic interval $(f_{x,y}, \rightarrow)$ and $(\leftarrow, f_{x,y})$ is open in the topology of pointwise convergence. This has already been done in Lemma 2 except for the cases $(\leftarrow, f_{x,y_0})$ and $(f_{x,y_1}, \rightarrow)$. Thus, we need only establish the following:

$$(\leftarrow, f_{x,y_0}) = \bigcup\{W(t, [y_0, y_1)): t < x\}, \tag{c}$$

$$(f_{x,y_1}, \rightarrow) = \bigcup\{W(t, (y_0, y_1]): t > x\}. \tag{d}$$

We shall prove (c). If $f_{z,w}<f_{x,y_0}$, then $z<x$ or $z=x$ and $w<y_0$. This latter case, however, is impossible. Since X is densely ordered, there exists $z<t<x$. Thus, $f_{z,w}(t)=y_0$, and $f_{z,w}\in W(t, [y_0, y_1))$. If $f_{z,w}\in W(t, [y_0, y_1))$ for some $t<x$, then $f_{z,w}(t)<y_1$. This implies that $z\leqq t<x$; so $f_{z,w}<f_{x,y}$ for all y. In particular, $f_{z,w}<f_{x,y_0}$. The proof of part (A) is now complete.

3. The Examples. An early example of a compact, perfectly normal, nonmetrizable space was given by Alexandroff and Urysohn [**1**, p. 76]. Their space, called $\boldsymbol{A}_7$, was constructed on the set of points of two half open intervals of real numbers one above the other, and can be described essentially as follows: Let $\boldsymbol{A}_7=[0, 1]\times\{0, 1\}-\{(0, 0), (1, 1)\}$. An arbitrary neighborhood $V(\xi)$ of a point $(\xi, 1)$ is the union of the half open interval $[\xi, x)\times\{1\}$ and the open interval $(\xi, x)\times\{0\}$. An arbitrary neighborhood $V(\eta)$ of a point $(\eta, 0)$ is the union of the half open interval $(y, \eta)\times\{0\}$ and the open interval $(y, \eta)\times\{1\}$. Alexandroff and Urysohn mentioned that $\boldsymbol{A}_7$ could be considered as an ordered space. Using the terminology of this note, we see that $\boldsymbol{A}_7$ can be obtained from $[0, 1]\times_{\text{lex}}\{0, 1\}$ by deleting the two isolated points $(0, 0)$ and $(1, 1)$ (in this case, the subspace topology is the same as the interval topology induced by the relative order). More recently, Ellis [**4**, p. 269] has given another construction of this space on a set consisting of two circles. It has also been noted by Johnson [**5**] that this space can be constructed on a single half open interval, or considered as the increasing functions from $[0, 1]$ to $\{0, 1\}$ (less the two constant functions).

A slight variation is given in Bourbaki [**2**, Section 2, Problem 13a, p. 49; Section 4, Problem 8, p. 97]: Let $E=[-1, 1]$. A base for the neighborhood system for each $x\in E$ is given by the set of all $U_\epsilon(x)=[x, x+\epsilon)\cup(-x-\epsilon, -x)$ for all $\epsilon>0$. This space can be obtained from $[0, 1]\times_{\text{lex}}\{0, 1\}$ by deleting the single isolated point $(0, 0)$.

4. Remarks on the Theorem. (1) In light of the statement of the above

theorem, it is interesting to note that $X \times_{\text{lex}} Y$ need not be homeomorphic to $X^{-1} \times_{\text{lex}} Y$. For example, let X be the set of ordinals less than the first uncountable ordinal, and let $Y=[0, 1)$ be a half open unit interval. Recall that $X \times_{\text{lex}} Y$, which is called the long line, is not metrizable. One can verify that $X^{-1} \times_{\text{lex}} Y$ is metrizable (it has a σ-locally finite base). If either X or Y admits an antiorder isomorphism onto itself (i.e., a map $g: X \rightarrow X$ which is 1-1 and onto and $x<x'$ is equivalent to $g(x')<g(x)$), then $X \times_{\text{lex}} Y$ and $X^{-1} \times_{\text{lex}} Y$ are order or antiorder isomorphic, and hence homeomorphic.

(2) In case $X=[0, 1]$ and $Y=\{0, 1\}$, then $\boldsymbol{F}$ consists of all the decreasing functions from X to Y. Likewise, the homeomorphic image of $[0, 1]^{-1} \times_{\text{lex}} \{0,1\}$ consists of all the increasing functions from $[0, 1]$ to $\{0, 1\}$.

(3) We can now show how our theorem can be used to prove that the examples (i) and (ii) are homeomorphic. The reason this is not immediate is that example (ii) refers to the increasing functions while part (A) of the theorem refers to the decreasing functions. In this case, however, with $X=[0, 1]$ and $Y=\{0, 1\}$ it follows from remark (1) that $X \times_{\text{lex}} Y$ is homeomorphic to $X^{-1} \times_{\text{lex}} Y$ which by remark (2) is homeomorphic to the increasing functions in Y^X.

(4) By use of methods similar to those used in the proof of our theorem, one can show that $\boldsymbol{F}$ (the closure of $\boldsymbol{F}$ in Y^X) is the largest superset of $\boldsymbol{F}$ which is linearly ordered by the product order. In fact, $\overline{\boldsymbol{F}}=\boldsymbol{F} \cup \{g \in Y^X : g(x)=y_0$ or y_1 for all $x \in X\}$. Finally, one can show that the order topology on $\overline{\boldsymbol{F}}$ agrees with the topology of pointwise convergence on $\overline{\boldsymbol{F}}$.

The author would like to thank P. R. Meyer and J. R. Hughes for references 1 and 4 respectively.

References

1. P. Alexandroff and P. Urysohn, Mémoire sur les espaces topologiques compacts, Verh. Nederl. Akad. Van Wetensch. Afd. Natuurk. Sect. 14, 1 (1929–1931) 1–96.
2. N. Bourbaki, Topologie Générale, Ch. 9, Hermann, Paris, 1958.
3. E. Čech, Topological Spaces, Interscience, New York, 1966.
4. R. Ellis and W. H. Gottschalk, Homomorphisms of transformation groups, Trans. Amer. Math. Soc., 94 (1960) 258–271.
5. R. A. Johnson, A compact nonmetrizable space such that every closed subset is a G-delta, this Monthly, 77 (1970) 172–176.
6. J. L. Kelley, General Topology, Van Nostrand, Princeton, 1955.
7. J. E. Vaughan, Embedding lexicographically ordered products in function spaces, Abstract 672-620, Notices, Amer. Math. Soc., 17 (1970) 262.

SUBSETS OF A COUNTABLE SET

J. R. Buddenhagen, Gogebic Community College, Ironwood, Michigan

The following problem is interesting in itself and has several applications (see, e.g., [1] and [2]). The usual solution involves rational valued sequences converging to irrationals. The author feels that the solution given here is more direct and better suited to classroom use.

Problem: Show that a countably infinite set has an uncountable family of subsets, the intersection of any two of which is finite.

Solution: Let L be the lattice points of E^2 (i.e., the points with integer coordinates). For each $\theta \in [0, \pi)$ let $S_\theta \subset L$ be the lattice points in an infinite strip with angle of inclination θ and width greater than 1. There are uncountably many S_θ, each infinite, but the intersection of any two is finite, since the intersection of any two of the strips is a bounded parallelogram in E^2.

References

1. P. R. Halmos, A Hilbert Space Problem Book, Van Nostrand, Princeton, N. J., 1967, pp. 5, 143, 170–171.

2. L. Gillman and M. Jerison, Rings of Continuous Functions, Van Nostrand. Princeton. N. J., 1960, pp. 79(5I), 97–98(6Q).

MATHEMATICAL EDUCATION

EDITED BY J. G. HARVEY AND M. W. POWNALL

Material for this Department should be sent to either of the editors: J. G. Harvey, Department of Mathematics, University of Wisconsin, Madison, WI 53706; M. W. Pownall, Department of Mathematics, Colgate University, Hamilton, NY 13346.

A TELEVISED COURSE FOR ELEMENTARY TEACHERS

J. E. LIGHTNER, Western Maryland College, and SISTER JOHN FRANCES GILMAN, Saint Joseph College, Emmitsburg, Maryland

Introduction. Mathematics education is currently in its second twentieth century "revolution," yet the need for in-service, fundamental mathematics courses for elementary school teachers still exists. To meet this need, in the spring of 1970 the Maryland School-College Mathematics Association, Inc. inaugurated a televised graduate credit mathematics course for elementary school teachers. The course was offered through the cooperation of four Maryland Colleges (Morgan State College, Towson State College, University of Maryland, and Western Maryland College), The Maryland Center for Public Broadcasting, and the Maryland State Department of Education-Division of Instructional Television. We wish here to discuss briefly the organization of the Maryland School-College Mathematics Association, Inc. (MS-CMA), and then to explain the planning, implementing and evaluation of the televised course.

History of MS-CMA. The Maryland School-College Mathematics Association was formed on March 10, 1964, and was incorporated in the state of Maryland on November 13, 1969. Its four-fold objectives are:

1. To promote greater cooperation and coordination between institutions concerned with the training of teachers of mathematics.
2. To stimulate greater articulation between elementary and secondary schools and institutions devoted to the preparation of teachers of mathematics.
3. To explore new directions and trends in the preparation of teachers.
4. To bring about an exchange of ideas and information for the improvement of mathematics instruction at all levels.

The association is composed of representatives of colleges and universities concerned with the preparation of teachers of mathematics, the State Department of Education, and the Mathematics Council of the Maryland State Teachers Association. By summer 1970, the membership included twenty-two colleges, universities, and other organizations in the State.

The first extensive project undertaken by MS-CMA was a series of conferences, conducted throughout the State, on the Mathematics Training of Elementary School Teachers, K-6. These conferences were conducted over a two year period and the final report was presented in the fall of 1967. This report indicated that each year the school systems of Maryland hire hundreds of elementary school teachers who bring a variety of collegiate backgrounds to the elementary classrooms of the state. In addition, the teachers who are already in the classrooms of the state have these same diverse backgrounds. Consequently, the elementary school teachers of the state have widely varying strengths and weaknesses. This same diversity is reflected when one focuses on a single area such as elementary school mathematics K-6. Among new college graduates there are some teachers who have had a strong mathematics program. On the average, however, elementary school teachers have had very little experience with mathematics courses in college or high school. Many teachers throughout the state report a fear or dislike of mathematics, but the attitude of the elementary school teacher toward mathematics is crucial, for too many children develop a fear of mathematics because their teachers are afraid of the subject. The report of the Conference also pointed out that many of these teachers expressed their need and desire for inservice-courses that will not only equip them with an in-depth understanding of contemporary mathematics but also with an enthusiastic interest in mathematics.

To meet this need, MS-CMA decided to consider the feasibility of a televised course in mathematics for elementary school teachers of grades K through 6. It was felt that television would make the course easily accessible to teachers who might not otherwise be able to go to a college campus or extension center for weekly sessions.

Cooperative planning. The MS-CMA first addressed itself to the possibilities of a televised course in the spring of 1967. By the spring of 1968, the MS-CMA established a committee to examine the existing mass-media mathematics courses for elementary school teachers of grades K through 6. The committee

was also asked to recommend a film series suitable from the viewpoint of content and methodology as well as suitable for telecasting throughout the State of Maryland. This committee consisted of representatives of those member colleges engaged in teacher education, elementary school teachers and supervisors, and teachers and supervisors who already had some experience with TV teaching.

The committee met during the summer of 1968 at the central office of the Washington County Board of Education in Hagerstown, Maryland. This county has for a number of years incorporated television as a part of its educational program. The committee, aided by the Coordinator of Mathematics of this county, viewed and evaluated a large number of existing television and film series dealing with mathematics for elementary school teachers. Evaluation was based on content, methodology, and suitability for state-wide telecasting. This committee suggested one series as suitable, and recommended further investigation of newer programs soon to be released. In the fall of 1968, MS-CMA established a second committee to move forward in the organization of a televised credit course for elementary school teachers of mathematics. The committee addressed itself to the three aspects of the total program:

1. *College Requirements.* The committee consulted with college presidents, academic deans, registrars, and financial officers regarding procedures and policies related to an inter-institutional program, and the willingness of these colleges to participate by granting credit, providing faculty, and providing facilities.
2. *School Needs.* Representatives of the State Department of Education were consulted regarding the feasibility of accepting for certificate renewal the graduate credit given by the colleges for the television course. Supervisors, teachers, and mathematics educators were consulted about the needs, expressed and unexpressed, of the Maryland elementary school teachers of mathematics, and about the adequacy of the film series proposed by the committee. On the basis of these discussions and further viewing, the final choice of the television series was made.
3. *Television Involvement.* The committee met with representatives of the Maryland Center for Public Broadcasting and the Maryland State Department of Education in order to discuss the technical aspects of the program which involved the television station, and to draw up a schedule of telecasts acceptable to the station and suitable to the elementary school teachers who would take the course. These representatives, all former teachers, worked with the committee to overcome many problems of scheduling and publicity.

The committee organized a one-semester course which employed twenty-eight videotapes and the accompanying worktext published by McGraw-Hill Book Company, *Modern Mathematics for the Elementary School Teacher*, by Kalin and Green. Two telelectures were shown each week, and every third week a two and one-half hour Saturday follow-up session was included as an integral part of the course. These follow-up sessions were designed to provide live instruction and discussion of mathematics, as well as interaction with fellow elementary school teachers and with the college instructor.

Many elementary school teachers, by their own admission, are fearful of mathematics. The Saturday sessions were intended to provide encouragement to these teachers by permitting them to recognize what they had learned

through television, as well as providing them with the opportunity to clarify points on which they may have been confused. The groups for the Saturday sessions were small (ranging from 13 to 28), thus increasing the benefits of instructor-teacher and teacher-teacher interaction. These follow-up sessions further provided each college with the opportunity to work directly with its own students, and offered the college faculty the opportunity to become aware of the needs of the elementary school teachers, and thus to plan appropriate future courses to meet these needs.

The course had for its principal audience approximately 260 elementary school teachers of mathematics in grades K through 6 who had minimal background in mathematics. Eighty-seven percent of these teachers already had the bachelor's degree and were qualified to take the course for graduate credit. The remainder of the enrollees did not yet have their bachelor's degree, and these latter took the course for undergraduate credit. No data are available on the number of persons who viewed the television series even though not formally registered for the course. About one-third of the participants had 4 or fewer years teaching experience, another one-third had 5–10 years experience, and the remainder had more than 10 years experience.

Organization and administration of the in-service course. The televised in-service course was offered under the cooperative sponsorship of four Maryland colleges, The Maryland Center for Public Broadcasting, the Maryland State Department of Education, and MS-CMA. Each participating college accepted the program as its own course, granted college credit of 3 semester hours to each person successfully completing the course (and meeting the other regular requirements for enrollment in that institution), determined the instructors who were involved in the Saturday follow-up sessions, and paid the instructors an appropriate salary. In addition, each college paid a share of the cost of the videotapes in proportion to its enrollment. The colleges provided six centers for the Saturday sessions; an attempt was made to locate these centers in a variety of geographical areas to meet the needs of those enrolled in the course. (MS-CMA had hoped to effect a plan by which teachers could take the course at the center nearest their residence and apply the credit to a degree at some other institution. Because of the variety of regulations for degrees at the different institutions, this plan proved unfeasible.)

Evaluation. The evaluation of the television course was based upon a final examination on the mathematical content of the course, and a questionnaire designed to obtain the opinions of the enrollees regarding the effectiveness of the telelectures, the television teachers, the worktext, and the Saturday sessions.

On the basis of the results of a course examination given in all centers on the same day, it was concluded that achievement in the course was high—the enrollees, on the whole, did better than satisfactory. An error analysis revealed certain areas of misunderstanding and confusion in the mathematical content. Based on this error analysis, the topics which the teachers in Maryland found

difficult were the same topics which Kalin and Green found to be areas of difficulty for a sample of teachers in Florida in 1967 in the pilot study of the televised course, *Modern Mathematics for Elementary Teachers*. The Maryland teachers, question for question, and over-all, performed essentially the same as those teachers in the pilot study. From this it was concluded that the course has been a satisfactory and successful one for the participants.

The participants were, in general, quite satisfied with the teachers in the television series, Robert Kalin and George F. Green, Jr., in regard to their ability to communicate their knowledge of subject matter, and their use of specially constructed visuals.

The majority of the participants viewed the telelectures in their homes, though many viewed them in their schools. In general, the participants were pleased with the arrangement of two different one-half hour telelectures each week. Each telelecture was broadcast four times each week, and a number of the participants watched some of the telelectures more than once. Most of the participants felt that the content covered in each lesson was appropriate, and that the pace of the course was reasonable.

The overwhelming majority of the enrollees felt that the worktext had been very helpful to them; in fact, the majority felt that in comparing the programmed text and the televised lectures, the worktext was more valuable to them. However, when they were asked whether they would prefer a televised course or a conventional course, they responded 2:1 in favor of the televised course. From this it was concluded that while the telelectures perhaps could be improved, in conjunction with the worktext and the follow-up sessions they proved to be an effective means of instruction.

The questionnaire further revealed that the follow-up sessions had been valuable for most of the teachers, and that their frequency (every third week) was appropriate to the needs of the enrolled teachers. A number of teachers, however, did feel that more follow-up sessions might be helpful.

In summary, the participants stated that the content of the course met their needs, was at a level appropriate to their mathematical background, and was presented in a novel and appealing way so as to motivate them to take other televised courses.

Summary. It is hoped that the success of this first venture will serve as an impetus for other graduate credit televised mathematics courses for inservice teachers (a second state-wide course is presently under consideration), and that colleges will construct new graduate programs, or a sequence of graduate courses, which emphasize mathematics for elementary school teachers. Ultimately, it is hoped that mathematically insecure, and poorly-prepared elementary school teachers will gain confidence, and will be prepared for enrollment in a graduate program leading to a master's degree.

At present, there are large numbers of elementary school teachers in Maryland who have little or no experience with mathematics courses in college.

This first course is designed to fill the basic mathematical needs of these teachers. In the future, after the basic needs are met, it is planned to expand the television course offerings so that Maryland teachers will be adequately prepared to meet the mathematics education needs of our school children.

PROBLEMS AND SOLUTIONS

EDITED BY EMORY P. STARKE

All problems (both elementary and advanced) proposed for inclusion in this Department should be sent to E. P. Starke, 1000 Kensington Ave., Plainfield, NJ 07060. Proposers of problems are urged to enclose any solutions or information that will assist the editors. Ordinarily, problems in well-known textbooks and results in generally accessible sources are not appropriate for this Department. No solutions (except those accompanying proposals) should be sent to Professor Starke.

ELEMENTARY PROBLEMS

Solutions of Elementary Problems should be sent to Problems Group, Mathematics Department, University of Maine, Orono, ME 04473. To facilitate their consideration, solutions of Elementary Problems in this issue should be typed (with double spacing) and should be mailed before August 31, 1971. Contributors (in the United States) who desire acknowledgment of receipt of their solutions are asked to enclose self-addressed stamped postcards.

E 2282 (1971, 196). **Correction.** We have received a correction on the statement of E 2282 from the proposer. In (2) the inequalities for the cases $\beta<60°$ and $\beta>60°$ should be interchanged to read:

$$0 < IH/IO < 1 \qquad \text{if } \beta > 60°$$

$$1 < IH/IO < 2 \qquad \text{if } \beta < 60°.$$

E 2295. *Proposed by R. S. Luthar, University of Wisconsin, Waukesha*

Suppose that m,n and $d(d>1)$ are arbitrary positive integers. Evaluate $(d^{md}-1,\ d^{nd}-1)$.

E 2296. *Proposed by Erwin Just and Norman Schaumberger, Bronx Community College*

A nonconstant polynomial f with integral coefficients has the property that for each prime p_i, there exists a prime q_i and an integer m_i such that $f(p_i)=q_i^{m_i}$. Prove that the polynomials contained in $\{x^n\}$, $n=1, 2, \cdots$, are the only polynomials which possess this property. (This generalizes E 1632 [1964, 795].)

E 2297. *Proposed by Richard Stanley, Harvard University*

Let $L(n)$ be the total number of distinct monomials appearing in the expansion of the determinant of an $n\times n$ symmetric matrix $A=(a_{ij})$. For instance, $L(3)=5$. Show that

$$\sum_{n=0}^{\infty} L(n)x^n/n! = (1-x)^{-1/2}\exp(\tfrac{1}{2}x+\tfrac{1}{4}x^2),$$

where $|x|<1$, and where we define $L(0)=1$.

E 2298. *Proposed by Anders Bager, Hjørring, Denmark*

Prove that in every triangle

$$\cos\frac{B-C}{2}+\cos\frac{C-A}{2}+\cos\frac{A-B}{2}$$
$$\leqq (\cos A+\cos B+\cos C)+\left(\sin\frac{A}{2}+\sin\frac{B}{2}+\sin\frac{C}{2}\right)\geqq 3,$$

with equality if and only if $A=B=C$.

E 2299. *Proposed by Anders Bager, Hjørring, Denmark*

It is given that the roots of a certain cubic equation

$$ax^3+bx^2+cx+d=0 \qquad (a\neq 0)$$

are $\tan(\tfrac{1}{4}A)$, $\tan(\tfrac{1}{4}B)$, and $\tan(\tfrac{1}{4}C)$, where A, B, C are the angles of a triangle. Prove that $a+b=c+d$.

E 2300. *Proposed by T. C. Brown, Simon Fraser University, British Columbia*

Let S be a semigroup in which, for some fixed $k\geqq 1$, $x^{k+1}=x$ and $xy^kx=yx^ky$ for all x, y in S. Show that S is commutative.

SOLUTIONS OF ELEMENTARY PROBLEMS

The Fermat Relation as a Matrix Equation

E 2030 [1967, 1133; 1968, 1123]. *Proposed by J. L. Brenner, University of Arizona, and Bernard Jacobson, Franklin and Marshall College.*

In the article, *Solutions of* $x^4+y^4=z^4$ in 2×2 *integral matrices* (this MONTHLY,

1966, p. 631) R. Z. Domaity gives examples. Show further (1) that $x^n+y^n=z^n$ is solvable in nonzero 2×2 integral matrices with nonnegative elements; (2) that $x^r+y^r=z^r$ is solvable in $r\times r$ nonsingular matrices with nonnegative elements.

Editorial Note. Some dissatisfaction has been expressed with the solution of (2) as originally presented. It seems unnatural to solve what purports to be analogous to Fermat's problem by the use of elements like $2^{1/n}I$. We call attention to a solution given by Bolker [1968, 759] which is described as follows:

Let $A(t)$ be the matrix (v_{ij}) in which $v_{12}=t$, $v_{n1}=1$, $v_{ij}=1$ for $2\leqq i\leqq n-1$, and $j=i+1$ with all other elements 0. Let a, b, c be positive integers satisfying $a+b=c$. Then the equation $X^n+Y^n=Z^n$ is satisfied by the nonsingular $n\times n$ matrices $X=A(a)$, $Y=A(b)$, $Z=A(c)$.

See also the additional note by P. M. Gibson in *Mathematics Magazine*, Nov. 1970, p. 275.

Permuting Paths

E 2240 [1970, 523]. *Proposed by Yeong-shyeong Tsai, Taiwan Provincial Chung-Hsing University*

Let n be an arbitrary positive integer, and let A be the set $\{1, 2, 3, \cdots, n\}$. Take n vertical line segments between two horizontal lines, indexing the top points and the bottom points by the numbers $1, 2, \cdots, n$. Now draw as many horizontal line segments as you please between adjacent vertical lines with the restriction that all endpoints of these segments are to be distinct. Now, starting at any upper point i, proceed downward until the first horizontal segment is reached, then go across that segment and again proceed downward; continue in this way, passing over each horizontal segment as it is met until a bottom point, say k_i, is reached.

(1) Prove that the mapping $i\rightarrow k_i (i=1, 2, \cdots, n)$ is a one-to-one mapping of A onto A.

(2) Prove that in completing the mapping $i\rightarrow k_i$ for all i as described above, every part of a vertical line is traced exactly once, and every horizontal segment is traced twice, once in each direction.

I. *Solution by Walter Bluger, Ottawa, Canada.* Assume that each element of A lies on one of the top points. At a given moment they all begin to move down at the same uniform speed. Hence two "neighbors" always arrive simultaneously at either endpoint of a horizontal segment. Then they both run across, in zero time, meeting each other at the midpoint, and continue descending until they all arrive simultaneously at the bottom line.

Each crossing produces only a reversal of neighbors and is thus traversed twice, and since all elements are at all times at the same height, each crossing is traversed twice only. Moreover, any vertical segment is travelled only once, because no element is ever higher than any other so that it could follow. The proof is thus evident.

II. *Solution by F. D. Parker, St. Lawrence University.* We shall call the horizontal line segments *rungs*. If there are no rungs, there is a one-to-one mapping (the identity mapping), each vertical line is traced exactly once, and each rung twice, once in each direction. If there is just one rung joining vertical lines i and j, the conclusions still hold, and the mapping can be described by the cycle (ij).

Suppose the result is true for m rungs, and that the one-to-one mapping is described by the m cycles of length two $(i_1j_1)\ (i_2j_2) \cdots (i_mj_m)$. We now add a new rung *lower* than any of the m rungs (we lose no generality in choosing this new rung to be the lowest). If this new rung joins vertical lines i_{m+1} and j_{m+1}, this new rung is traced twice, once in each direction, and the mapping is now described by the cycles $(i_1j_1)\ (i_2j_2) \cdots (i_mj_m)\ (i_{m+1}j_{m+1})$.

Also solved by J. D. Baum, D. G. Beane & E. F. Schmeichel, Roger Giudici, Michael Goldberg, G. A. Heuer & C. V. Heuer, F. W. Humburg, David Kelly, Harry Lass, J. F. Leetch, J. V. Michalowicz, Norman Miller, D. E. Penney, J. Pfaendtner, Simeon Reich (Israel), David Singmaster (England), R. F. Smith, and the proposer.

Hobson's Choice Integral

E 2241 [1970, 523]. *Proposed by Linda Pleska, Bowling Green State University*

Prove or disprove: If $\lim_{x\to x_0} f(x)$ exists for each $x_0 \in [a, b]$, then the Riemann integral $\int_a^b f(x)dx$ exists.

Editorial Note. The editors acknowledge with embarrassment that this problem is known almost everywhere. Forty-one readers offered solutions or provided references from Hobson to the *Mathematics Magazine* (1967, 199; 232) and this MONTHLY (1970, 412).

An Exponential Sum Modulo n

E 2242 [1970, 652]. *Proposed by M. L. Fredman, California Institute of Technology*

Show that $\sum_{i=1}^n k^{(i,n)} \equiv 0 \pmod n$, $n \geqq 1$, for all integers k (positive or negative), where (a, b) denotes the greatest common divisor.

I. *Solution by L. E. Mattics, University of South Alabama.* We have

$$\sum_{i=1}^{n} k^{(i,n)} = \sum_{d|n} \phi\left(\frac{n}{d}\right)k^d,$$

and, if m and n are relatively prime, this gives

$$\sum_{i=1}^{n\cdot m} k^{(i,m\cdot n)} = \sum_{d|m} \phi\left(\frac{m}{d}\right) \sum_{\tau|n} \phi\left(\frac{n}{\tau}\right)(k^d)^\tau.$$

Hence it is clear that we need prove the conjecture only for $n = p^\alpha$ where p is prime and $\alpha \geqq 1$.

The cases $\alpha = 1$ or $p|k$ are trivial, so let $\alpha > 1$ and $(k, p) = 1$. Then

$$(*)\qquad \sum_{i=0}^{\alpha}\phi\left(\frac{p^\alpha}{p^i}\right)k^{p^i} \equiv k^{p^\alpha}+\phi(p)k^{p^{\alpha-1}}+\phi(p^2)k^{p^{\alpha-2}}+\cdots+\phi(p^{\alpha-1})k^p+\phi(p^\alpha)k$$
$$\equiv k^{p^{\alpha-1}}+(p-1)k^{p^{\alpha-1}}+\cdots+\phi(p^\alpha)k$$
$$\equiv p(k^{p^{\alpha-1}}+\cdots+\phi(p^{\alpha-1})k)$$
$$\equiv p\sum_{i=0}^{\alpha-1}\phi\left(\frac{p^{\alpha-1}}{p^i}\right)k^{p^i} \pmod{p^\alpha}$$

and we are done by induction. Note the use of Fermat's theorem in the reduction of the first term in the right hand side of (*).

II. *Solution by L. Carlitz, Duke University.* Put

$$S(n,k)=\sum_{i=1}^{n}k^{(i,n)},\qquad T(n,k)=\sum_{ab=n}\mu(a)k^b.$$

Then

$$S(n,k)=\sum_{d\mid n}k^d\sum_{\substack{s=1\\(s,n/d)=1}}^{n/d}1=\sum_{d\mid n}\phi(n/d)k^d,$$

$$\sum_{ab=n}aT(b,k)=\sum_{abc=n}a\mu(b)k^c=\sum_{mc=n}\phi(m)k^c,$$

so that

$$(1)\qquad S(n,k)=\sum_{ab=n}aT(b,k).$$

Now it is well known that

$$(2)\qquad T(n,k)\equiv 0 \pmod{n}.$$

(See, e.g., Dickson, *History of the Theory of Numbers*, vol. 1, p. 84.) It follows at once from (1) and (2) that $S(n,k)\equiv 0 \pmod{n}$.

Also solved by D. M. Bloom, Neal Felsinger, M. G. Greening (Australia), J. J. Heed, V. S. Joshi & A. M. Vaidya (India), Harry Lass, Gustav Lehrer (Norway), Andrzej Makowski (Poland), David Monk (Scotland), Bob Prielipp, Simeon Reich (Israel), St. Olaf College Students, E. F. Schmeichel, David Singmaster (England), Allen Stenger, Walter Stromquist, E. W. Trost (Switzerland), L. J. Warren, and the proposer.

Note. Several solvers point out that

$$\frac{1}{n}\sum_{i=1}^{n}k^{(i,n)}=\frac{1}{n}\sum_{d\mid n}\phi\left(\frac{n}{d}\right)k^d$$

is the number of circular permutations of k distinct elements taken n at a time (repetitions allowed), as was proved by P. A. MacMahon in Proc. London Math. Soc., 23 (1891-2) 305–313. Singmaster notes that the problem is mentioned in his paper, *A Maximal generalization of Fermat's Theorem;* Mathematics Magazine, 39, March 1966, 103–107.

Doubly Stochastic Matrices

E 2243 [1970, 652]. *Proposed by Richard Sinkhorn, University of Houston*

Show that every normal stochastic matrix is necessarily doubly stochastic.

Solution by R. C. Thompson, University of California at Santa Barbara. The solution is a special case of the following more general result:

THEOREM. *If A is a normal matrix with real or complex entries and if each row sum of A is c, then each column sum of A also equals c.*

Proof. Let A be an $n\times n$ matrix which satisfies the hypotheses. Because A is normal, $A^*=\overline{A}^t$ is a polynomial in A, say $A^*=p(A)$. Let J be the $n\times n$ matrix in which each entry is 1. $AJ=cJ$. Hence $A^tJ=c^tJ$ and $A^*J=p(A)J=p(c)J$. It now follows that $A^tJ=\overline{p(c)}J$. Therefore each column sum of A equals $\overline{p(c)}$. Sum all the elements of A in two ways, first by rows and then by columns. By equating the results it follows that $nc=n\overline{p(c)}$. Therefore $\overline{p(c)}=c$ and $A^tJ=cJ$. It now follows that $JA=cJ$ and consequently each column sum of A is equal to c as claimed. The proposed problem is the case $c=1$.

Also solved by D. J. Bordelon, Emeric Deutsch, J. W. Duke, I. V. Gentry, M. G. Greening (Australia), W. E. Hoff, C. R. Johnson, Harry Lass, Gustav Lehrer (Norway), Joel Levy, Marvin Marcus, J. P. McLean, Henryk Minc, David Monk (Scotland), P. J. Nikola, K. R. Rebman, Simeon Reich (Israel), Sid Spital, Olga Taussky Todd, J. R. Ventura, Jr., E. T.-H. Wang, and the proposer.

Editor's Note. O. T. Todd pointed out that she communicated a solution of this problem to Miss E. Haynsworth in 1953. Miss Haynsworth published a generalization of this result in *Quasi-stochastic matrices*, Duke Mathematical Journal, 1955.

Sum of a Series

E 2244 [1970, 652]. *Proposed by T. K. Leong, T. A. Peng and K. C. Yeo, University of Singapore*

Show that, for any fixed $m\geqq 2$, the series

$$1+\frac{1}{2}+\cdots+\frac{1}{m-1}-\frac{x}{m}+\frac{1}{m+1}+\frac{1}{m+2}+\cdots+\frac{1}{2m-1}-\frac{x}{2m}$$
$$+\frac{1}{2m+1}+\frac{1}{2m+2}+\cdots+\frac{1}{3m-1}-\frac{x}{3m}+\cdots$$

is convergent for exactly one value of x and find the sum of the series for this x.

Solution by T. J. Cullen, California State Polytechnic College, Pomona. Let

$$S_n(x)=\sum_{k=1}^{n}\left[\frac{1}{(k-1)m+1}+\cdots+\frac{1}{km-1}-\frac{x}{km}\right].$$

If the given series converges for x and y, then the sequence

$$S_n(x)-S_n(y)=\frac{y-x}{m}\sum_{k=1}^{n}\frac{1}{k}$$

converges. But this is only possible if $x=y$, since $\sum 1/k$ diverges. Hence the series converges for at most one value of x.

The sequence $A_n=1+\frac{1}{2}+\cdots+1/nm-\log(nm)$ is known to converge to Euler's constant γ. Now

$$S_n(m-1) = A_n + \log(nm) - \sum_{k=1}^{n}\frac{1}{km} - \sum_{k=1}^{n}\frac{m-1}{km}$$

$$= A_n + \log m + \left(\log n - \sum_{k=1}^{n}\frac{1}{k}\right) \to \gamma + \log m - \gamma = \log m.$$

Therefore, when $x=m-1$, the series converges to $\log m$.

Also solved by J. A. Belward (Australia), M. T. Bird, J. R. Blake (Australia), D. M. Bloom, R. H. Brown, T. C. Brown, L. Carlitz, M. S. Demos, W. E. Dydo, R. L. Evison, Charles Fox, Michael Goldberg, M. T. Greening (Australia), Emil Grosswald, Judith R. Gumerman, J. J. Heed, Robert Heller, G. A. Heuer, G. A. & C. V. Heuer, Dean Hickerson, Douglas Holdridge, R. T. Hood, L. N. Howard, P. G. Kirmser, Christopher Landauer & Robert Patenaude, O. P. Lossers (Netherlands), M. E. Low, Henrik Meyer (Denmark), Norman Miller, M. H. Moore, M. L. Mumford, D. E. Penney, M. A. Radke, Simeon Reich (Israel), Eric Rosenthal, Bert Ross, St. Olaf College Students, A. A. Sardinas, E. F. Schmeichel, Paul Schweitzer, Sid Spital, Walter Stromquist, A. M. Vaidya & H. N. Rawal & C. G. Khatri (India), D. J. Wagner, L. E. Ward, Sr., K. L. Yocom, and the proposers.

ADVANCED PROBLEMS

All solutions of Advanced Problems should be sent to J. Barlaz, Rutgers—The State University, New Brunswick, N.J. 08903. Solutions of Advanced Problems in this issue should be typed (with double spacing) on separate, signed sheets and should be mailed before August 31, 1971. Contributors (in the United States) who desire acknowledgment of receipt of their solutions are asked to enclose self-addressed stamped postcards.

5795. *Proposed by B. R. Myers, University of Notre Dame*

An n-wheel is a graph consisting of one "outer" circuit having n vertices and edges along with the n edges connecting these vertices to a single "hub" vertex. A spanning tree of a graph on $(n+1)$ vertices is a collection of n edges in the graph which contains no circuit.

How many different spanning trees are there in an n-wheel? (The result is conveniently expressible in terms of Fibonacci numbers.)

5796.* *Proposed by R. S. Luthar, University of Wisconsin at Waukesha*

Show that, $\phi(n)$ being the Euler totient,

$$\limsup_{n\to\infty} \frac{\phi(n+1)}{\phi(n)} = \infty, \qquad \liminf_{n\to\infty} \frac{\phi(n+1)}{\phi(n)} = 0.$$

'5797. *Proposed by I. N. Herstein, University of Chicago, and Susan Montgomery, University of Southern California*

A theorem of Marshall Osborn states: If R is a simple ring of characteristic not 2 and with an involution such that every nonzero symmetric element is invertible, then either R is a division ring or is 4-dimensional over its center. Show that if R is a prime ring with involution, of characteristic 2 and if every nonzero symmetric element of R is invertible, then R must be a division ring. (*Prime* means $xRy=0$ implies $x=0$ or $y=0$.)

5798. *Proposed by C. J. Eliezer, La Trobe University, Bundoora, Australia*

Prove that for $x>1$ and $y>1$,

$$\frac{\Gamma(x)}{(x-1)^x} + \frac{\Gamma(y)}{(y-1)^y} \geqq \frac{2\Gamma\left(\dfrac{x+y}{2}\right)}{\left(\dfrac{x+y}{2}-1\right)^{x+y}},$$

$$\frac{\Gamma(x)\Gamma(y)}{\left[\Gamma\left(\dfrac{x+y}{2}\right)\right]^2} \geqq \frac{(x-1)^x(y-1)^y}{\left(\dfrac{x+y}{2}-1\right)^{x+y}}.$$

5799. *Proposed by C. J. Eliezer, La Trobe University, Australia*

Prove that for $-1<p<1$,

$$\frac{1}{p+1} - \frac{1}{p+2} + \frac{1}{p+3} - \cdots \geqq \frac{1-4p+2p^2}{(1-p)(2-p)}, \quad \text{and}$$

$$\frac{1}{p+1} - \frac{1}{p+2} + \frac{1}{p+3} - \cdots \geqq \frac{(1-p)(2-p)}{(3-2p)}.$$

5800. *Proposed by Joel Pitcairn, Huntingdon Valley, Pa.*

Exercise 16.1 of Halmos, *Measure Theory* says: If E is a Lebesgue measurable set such that, for every x in a dense set, $\mu(E\Delta(E+x))=0$, then $\mu(E)=0$ or $\mu(E')=0$. Prove the following generalization (which is useful for producing 'maximally nonmeasurable' sets): If E is a subset of a locally compact group (with left Haar measure μ) such that, for every x in a dense set, $E\Delta xE$ is locally null, then either (1) E is locally null or (2) E' is locally null or (3) $\mu^*(A\cap E) = \mu^*(A\cap E') = \mu(A)$ for every measurable set A. [A set is locally null if its intersection with every compact set has measure 0.]

5801. *Proposed by Erwin Just, Bronx Community College*

If m and k are arbitrary fixed positive integers and m is odd, prove that: (1) There exists a positive integer n such that m^n+n^m contains at least k distinct prime factors, and (2) there exists a positive integer t such that $m^{t+j}+(t+j)^m$ is composite if $j\in\{1, 2, \cdots, k\}$.

SOLUTIONS OF ADVANCED PROBLEMS

Compactification of Metric Spaces

5725 [1970, 313]. *Proposed by D. J. Lutzer, University of Washington*

Is it true that every compact Hausdorff space is a compactification of a metric space?

Solution by J. W. Taylor, University of Illinois. No. In fact, the product X of uncountably many copies of the closed unit interval of the real line is not the compactification of a first countable space, nor even of a space which is first countable at least at one point. It is known and easily proved that if D is a dense subspace of a regular space Y and N is a neighborhood base in D of a point d, then $\{\text{closure}_Y\ V: V\in N\}$ is a neighborhood base for d in Y. Thus a dense subspace of X cannot be first countable at any point, for then X would be first countable at that point; but X is not first countable at any of its points.

Also solved by Cleveland State University Problem Group, A. A. Jagers (Netherlands), M. R. Kirch, H. E. Lacey, Warren Page, and the proposer.

Cartesian Product Measures

5729 [1970, 410]. *Proposed by L. F. Kemp, Jr., Polytechnic Institute of Brooklyn*

Let $u\times v$ and $u'\times v'$ be two Cartesian product measures defined on $(X\times Y, S\times T)$, the Cartesian product of two measure spaces (X, S) and (Y, T). Then

1. $u \times v \ll u' \times v'$ if and only if $u \ll u'$ and $v \ll v'$.
2. $u \times v \perp u' \times v'$ if and only if $u \perp u'$ or $v \perp v'$.

Solution by John Milcetich, Federal City College, Washington, D. C.

1. We assume that neither u nor v is the zero measure, for otherwise the statement would not follow.

Suppose $u\times v\ll u'\times v'$ and $u'(E)=0$ for some $E\in S$. Since v is not the zero measure, there is $F\in T$ such that $v(F)\neq 0$. Then $u'\times v'(E\times F)=u'(E)\cdot v'(F)=0$. Therefore $0=u\times v(E\times F)=u(E)\cdot v(F)$. Hence $u(E)=0$ and $u\ll u'$. Similarly $v\ll v'$.

Suppose $u\ll u'$ and $v\ll v'$. Let $u'\times v'(H)=0$. Then $v'(H_x)=0$ for all $x\in X-E$, where $u'(E)=0$. Since $v\ll v'$ and $u\ll u'$, $v(H_x)=0$ for all $x\in X-E$ and $u(E)=0$. Then

$$u \times v(H) = \int_X v(H_x)du = \int_{X-E} v(H_x)du = 0.$$

Hence $u\times v \ll u'\times v'$.

2. Suppose $u\perp u'$. There are disjoint nonempty E, $F\in S$ such that $u(E)=0=u'(F)$ and $E\cup F=X$. Then $E\times Y$, $F\times Y$ are disjoint, nonempty $u\times v$-measurable sets: $E\times Y\cup F\times Y=X\times Y$ and $u\times v(E\times Y)=0$ and $u'\times v'(F\times Y)=0$. Hence $u\times v\perp u'\times v'$. Similarly $v\perp v'$ implies $u\times v\perp u'\times v'$.

If $u\times v\perp u'\times v'$, then there are distinct, nonempty A, $B\in S\times T$ such that $A\cup B=X\times Y$ and $u\times v(A)=0=u'\times v'(B)$. Let $E=\{x\in X;\ v(A_x)\neq 0\}$ and $E'=\{x\in X: v'(B_x)\neq 0\}$. Then $u(E)=0=u'(E)$. If $E\cup E'=X$, let $F\subset E$, $F'\subset E'$ such that F and F' are nonempty and disjoint and $F\cup F'=X$. Then $u(F)=0=u'(F')$, so $u\perp u'$. If $E\cup E'\neq X$, there is $x\in X$ such that $v(A_x)=0=v'(B_x)$. Then we find disjoint, nonempty F and F', $F\subset A_x$, $F'\subset B_x$ such that $F\cup F'=Y$. Then $v(F)=0=v'(F')$ and $v\perp v'$.

Also solved by D. A. Hejhal, and by the proposer.

Non-Topological Groups

5732 [1970, 410]. *Proposed by James Chew, University of Akron, Ohio*

Prove or disprove: Let $(X,\ \cdot,\ \mathfrak{J})$ be a system such that $(X,\ \cdot)$ is a group and $(X,\ \mathfrak{J})$ is a topological space such that multiplication is continuous. If $\text{card}(W\cap W^{-1})\geqq 2$ for every open set W containing the identity, then $(X,\ \cdot,\ \mathfrak{J})$ is a topological group.

Solution by Bob Gray, Los Alamos Scientific Laboratory. The proposition is false. Consider the real numbers R under addition and topologized with the right-hand topology, that is, the topology with base sets of the form (a,∞). Multiplication is continuous: let a, $b\in R$; now any neighborhood of $a+b$ contains a set of the form $(a+b-\epsilon,\ \infty)$, and $(a-\epsilon/2,\ \infty)$ and $(b-\epsilon/2,\ \infty)$ are neighborhoods of a, b, respectively, such that $(a-\epsilon/2,\ \infty)+(b-\epsilon/2,\ \infty)\subset(a+b-\epsilon,\ ,\infty)$. If U is a neighborhood of 0, it contains $(-\epsilon,\infty)$ for some ϵ and $U\cap U^{-1}\supset(-\epsilon,\ \epsilon)$. However, inverse is not continuous: let $a\in R$, then $(a-1,\ \infty)$ is a neighborhood of a, but neighborhoods of $-a$ contain a set of the form $(-a-\epsilon,\infty)$, and $-(-a-\epsilon,\ \infty)=(-\infty,a+\epsilon)$ is not contained in $(a-1,\infty)$.

Note that the right-hand topology is T_0 but not T_1. The proposition still remains false if we require the topology to be Hausdorff. This can be seen by considering the product formed by the reals under addition with the usual topology and the reals under addition with the upper-limit topology, that is, the topology whose base sets are of the form $(a,\ b]$.

Also solved by D. E. Beken, S. L. Campbell, Helen F. Cullen, D. L. Grant, A. A. Jagers (Netherlands), J. O. Kiltinen, Jürg Rätz (Switzerland), and L. E. Ward, Jr.

Independent Random Variables

5733 [1970, 410]. *Proposed by M. F. Neuts, Purdue University*

Let X be a nonnegative integer-valued random variable, and suppose that

Y is a random variable satisfying $0 \leqq Y \leqq X$. We are interested in the property ($\mathcal{I}$): Y and $X-Y$ are independent.

(A) If Y takes only integer values, and if the conditional distribution of Y, given the value of X, is uniform on $\{0, 1, \cdots, X\}$, show that ($\mathcal{I}$) holds if and only if Y and $X-Y$ have geometric distributions with the same parameter.

(B) If Y takes real values and has uniform conditional distribution, given X on the interval $0 \leqq Y \leqq X$, when does ($\mathcal{I}$) hold?

Solution by Ellen Hertz, Columbia University. (A) Let $p_n = P(X=n)$, $q_n = P(Y=n)$, $n = 0, 1, 2, \cdots$. Then $q_k = \sum_{n=k}^{\infty} p_n/(n+1)$ and

$$P(X - Y = k) = \sum_{n=k}^{\infty} P(X = n \quad \text{and} \quad Y = n - k) = \sum_{n=k}^{\infty} p_n/(n + 1) = q_k.$$

Then $P(Y=j \text{ and } X-Y=m) = P(Y=j \text{ and } X=m+j) = p_{m+j}/(m+j+1) = q_{m+j} - q_{m+j+1}$. Then $X-Y$ and Y independent implies

$$q_{m+j} - q_{m+j+1} = q_j q_m. \tag{1}$$

Then ($\mathcal{I}$) implies $q_{m+1} = q_m(1-q_0)$, so that

$$q_n = q_0(1 - q_0)^n, \qquad n = 0, 1, 2, \cdots. \tag{2}$$

But (2) implies (1), so that ($\mathcal{I}$), independence of $X-Y$ and Y, is equivalent to (2).

B. ($\mathcal{I}$) holds if and only if Y and $X-Y$ are exponential with the same parameter.

Proof: Let $F(x) = P(X \leqq x)$, $H(y) = P(Y \leqq y)$. Then

$$1 - H(y) = \int_{x=y}^{\infty} [(x - y)/x] dF(x) \qquad (y \geqq 0)$$

and

$$P(X - Y \geqq y) = P(Y \leqq X - Y) = \int_{x-y \geqq 0} [(x - y)/x] dF(x) = 1 - H(y).$$

Also,

$$\begin{aligned} P(Y \geqq y \text{ and } X - Y \geqq z) &= P(Y \geqq y \text{ and } Y \leqq X - z) \\ &= \int_{x=y+z}^{\infty} [(x - z - y)/x] dF(x) = 1 - H(y + z). \end{aligned} \tag{3}$$

According to (3), ($\mathcal{I}$) is equivalent to

$$1 - H(y + z) = (1 - H(y))(1 - H(z)) \qquad (y, z \geqq 0). \tag{4}$$

But since $1-H$ is bounded and not identically zero [Feller, *An Introduction to Probability Theory and its Applications*, Vol. 1, p. 413] it follows that $1-H(x) = e^{-cx}$ for some constant c.

Also solved by L. E. Clarke (England), J. C. Hickman, David Kelly, Frank Knight, the proposer, and one who omitted his signature.

Editorial Note. The solution of (B) is not based on X taking only integer values.

Proximity of Bernstein Polynomials to x^n

5734 [1970, 532]. *Proposed by G. G. Lorentz, University of Texas*

Let E_n be the degree of uniform approximation of x^{n+1} on the interval $[0, 1]$ by polynomials

$$(\dagger) \qquad P_n(x) = \sum_{k=0}^{n} a_k x^k (1-x)^{n-k}, \qquad a_k \geqq 0$$

with positive coefficients in $x(1-x)$. In other words, let E_n be the minimum of $\max_{0\leqq x\leqq 1}|x^{n+1}-P_n(x)|$ for all polynomials of the form (†). Prove that $y=\lim_{n\to\infty} nE_n$ exists, and find y.

Solution by C. S. Gardner, University of Texas. The best fit is given by $P_n=ax^n$ where a and x_1, $0<a<1$, are determined by the relations:

$$(1) \qquad a x_1^n - x_1^{n+1} = 1 - a, \qquad x_1 = na/(n+1).$$

For if

$$P_n^* = \sum_{k=0}^{n} a_k^* x^k (1-x)^{n-k},$$

then either $|1-P_n^*(1)| = |1-a_n^*| \geqq 1-a$; or else $a_n^* \geqq a$, in which case

$$P_n^*(x_1) - x_1^{n+1} = a x_1^n - x_1^{n+1} = 1 - a,$$

and thus the assertion is proved.

Furthermore, since

$$\frac{d}{dx}(ax^n - x^{n+1}) = [na - (n+1)x]x^{n-1}$$

is positive for $x<x_1$ and negative for $x>x_1$, it follows that

$$\max|ax^n - x^{n+1}| = 1 - a = E_n.$$

Setting $E_n = y_n/n$ we obtain

$$[n/(n+1)]^{n+1}(1 - y_n/n)^{n+1} = y_n,$$

whence $0<y_n<1$. Hence as $n\to\infty$, $\overline{\lim}\, y_n = z$ and $\underline{\lim}\, y_n = w$ are both in $[0, 1]$. Picking a subsequence of $\{y_n\}$ which tends to z we get in the limit $e^{-1}e^{-z}=z$, that is $ze^z=e^{-1}$. Likewise $we^w=e^{-1}$. Since xe^x is montone increasing $(x>0)$ we have $z=w=y$ and hence $\lim nE_n = y$, where y is the positive root of $ye^y=e^{-1}$.

An r-th Order Nonlinear Difference Equation

5736 [1970; 532, 774]. *Proposed by M. S. Klamkin, Ford Scientific Laboratory*

Solve the nonlinear difference equation of the rth order

$$D_n = a_1 D_{n-1}^{m+1} + a_2 D_{n-1}^{m} D_{n-2}^{m+1} + \cdots + a_r D_{n-1}^{m} D_{n-2}^{m} \cdots D_{n-r+1}^{m} D_{n-r}^{m+1},$$

(m, r, a_i, constants).

Solution by the proposer. By considering the case $r=2$, one is led, after some trial and error, to rewrite the given equation in the form

$$1 = a_1\phi_n + a_2\phi_n\phi_{n-1} + \cdots + a_r\phi_n\phi_{n-1} \cdots \phi_{n-r+1},$$

in which we have replaced D_{n-1}^{m+1}/D_n by ϕ_n. By letting $\phi_n = \psi_n/\psi_{n+1}$, we obtain the linear difference equation

$$\psi_{n+1} = a_1\psi_n + a_2\psi_{n-1} + \cdots + a_r\psi_{n-r+1},$$

which has the general solution

$$\psi_n = \sum_{i=1}^{r} k_i R_i^n,$$

where R_i are the roots of $x^r = a_1 x^{r-1} + a_2 x^{r-2} + \cdots + a_r$.

Retracing our substitutions, we get in turn

$$\phi_n = \psi_n/\psi_{n+1}, \qquad D_n\phi_n = D_{n-1}^{m+1},$$

or equivalently, $\log D_n = (m+1)\log D_{n-1} - \log \phi_n$. Let $\log D_n = (m+1)^n A_n$; then $A_n - A_{n-1} = -\log \phi_n^{(m+1)^{-n}}$. Thus

$$A_n = -\log\left\{e^{-A_0}\prod_{j=1}^{n}\phi_j^{(m+1)^{-j}}\right\},$$

and finally,

$$D_n = e^{A_0(m+1)^n}\prod_{j=1}^{n}\phi_j^{-(m+1)^{n-j}}.$$

The equation arose as a generalization in a study of the frequency spectrum of a mass-spring system which forms a rooted Cayley tree.

Also solved by L. Carlitz.

REVIEWS

EDITED BY J. ARTHUR SEEBACH, JR. AND LYNN A. STEEN
with the assistance of the mathematics departments of St. Olaf and Carleton Colleges

COLLABORATING EDITOR FOR FILMS: SEYMOUR SCHUSTER, Carleton College

Printed materials for review should be sent to: Book Review Editor, American Mathematical Monthly, St. Olaf College, Northfield, MN 55057. Films and correspondence relating to films should be sent to Seymour Schuster, Carleton College, Northfield, MN 55057.

All unsigned material is written by one of the editors. A boldface capital C in the margin indicates that a review is based in part on classroom use. Professors willing to write such a review should inform the editor in order to avoid duplication.

Hilbert. By Constance Reid. Springer Verlag, New York—Heidelberg—Berlin, 1970. xi+290 pp. Frontispiece and 28 illustrations, $8.80. (Telegraphic Review, June–July 1970.)

For the older mathematician this is a captivating, for the younger mathematician an important, book—because of both its subject and its genre. Also in non-mathematical circles, it deserves and undoubtedly will receive widespread reading. As a biography of David Hilbert, it is not just an account in tombstone marble, but a portrait with color, light, and atmosphere. Beyond this, it preserves—while still in the memory of living men– a real sense of the style and human identity that characterizes a significant era in the development of mathematics.

The life of Hilbert was inextricably interwoven with mathematics at Göttingen—the golden epoch, presided over by Minkowski, Hilbert, and Klein before the first World War; and the more agitated, tougher, but very productive years between the two wars which had finally come to an end with the historical reply of the frail old man to a high Nazi functionary: "Mathematics at Göttingen? There really isn't any more." We know the tell-tale anecdotes and legends of those times: with the indefatigability of a good chronicler, Miss Reid has put together the fabric of fact to which they belong. Here are the well-known personalities and the radiance of their imaginative work. Here also, caught almost inadvertently by the chronicler's honesty, are their typical inhumanities, conditioned by the age and their status.

For the mathematician, enough is said about the mathematics that occupies Hilbert and that happens around him to remind and evoke. Full measure on the significance and the impact of Hilbert's mathematical work, up to the time of his death, is provided by reprinting—without the purely biographical paragraphs—Herman Weyl's *David Hilbert and His Mathematical Work* (Bull. Amer. Math. Soc., 50 (1944) 612–654) at the end of the book. For the non-mathematician, the treatment is light enough to leave him, but rarely and then for short stretches, beyond the line of understanding. It is perhaps a drawback that it allots less than the space they deserve to Hilbert's disciples in mathematics and more than necessary to his "tutors in physics." In summary, this account is not, nor pretends to be, a definitive science-historical appraisal, but it is a valuable

record without the publication of which much information about David Hilbert and his time could have been lost for good.

F. J. WEYL, Hunter College

Single Variable Calculus, with an Introduction to Numerical Methods. By Melvin Henriksen and Milton Lees. Worth, New York, 1970. xv+624 pp. $10.95. (Telegraphic Review, May 1970.)

Henriksen and Lees have written an interesting, stimulating text which this reviewer is glad to have read and will refer to because of its novelty, excellent exercises, and generally careful style. As the title of the book indicates, numerical methods are emphasized, and they are handled with care. Every computation is accompanied by an error analysis (which sometimes makes for heavy going in a text addressed to freshmen). The author's emphasis on numerical computations and on estimation of errors is good mathematics, good pedagogy, and useful to most students. A by-product of the emphasis on estimation of errors is extensive, meaningful practice with absolute values, inequalities, and algebraic manipulation.

The derivative is introduced by means of admissible null sequences. If x is in the domain of a function f, then $\langle h_n \rangle$ is an admissible null sequence if $h_n \neq 0$, and $x+h_n$ is in the domain of f for every positive integer n. Function f has derivative m at x if $m-[f(x+h_n)-f(x)]/h_n$ is a null sequence whenever $\langle h_n \rangle$ is an admissible null sequence. One-sided derivatives at the end of any interval in the domain of f are automatically included. Continuity is defined similarly. The limit of a function of a continuous variable is first introduced on page 427 in connection with integrals improper at infinity. The limit from the right is then defined by $\lim_{x \to a+} f(x) = \lim_{x \to \infty} f(a+1/x)$. The limit exists if the one-sided limits are equal.

The antiderivative is introduced as a solution of the initial value problem $\phi' = f$, $\phi(a) = p$. The development is done well. One of the interesting and somewhat novel examples is determination of the antiderivative of $|x|$. The exponential and sine functions are defined as the unique solutions of the initial value problems $\phi' = \phi$, $\phi(0) = 1$, and $\phi'' + \phi = 0$, $\phi(0) = 0$, $\phi'(0) = 1$. As they do with most topics, the authors develop the properties of these functions carefully.

Topics are well-motivated by applications and by the needs of the mathe matics itself. There are many nice touches. Instead of the usual x and y axes, there are first and second coordinates. Difficulties arising from the symbolism "$y = f(x)$" are avoided by never using it. The proofs of the properties of the absolute value function are among the neatest this reviewer has seen. To obtain the derivative of x^n, the authors determine the derivative of $xf(x)$, and thereby motivate the idea of continuity.

Although careful attention is paid to graphing, there is no treatment at all of analytic geometry. The conic sections are not even mentioned. Perhaps this omission is unwise, because of the applications of the conics in physics. Geom-

etry, particularly those parts of it which are intimately associated with linear algebra and analysis, should not be neglected.

Proofs of any depth are relegated to a sequence of problems in an appendix. They are presented well with ample hints. The relatively easy proof of the existence of the Riemann integral for continuously differentiable functions is included. The proof of the existence of the integral for continuous functions is given in a later appendix. A sequel, "Several Variable Calculus" by Milton Lees, will be published in 1971.

SYLVAN WALLACH, C. W. Post College

Modern General Topology. By Jun-iti Nagata. Bibliotheca Mathematica, Vol. 7. North-Holland, Amsterdam, 1968. Distributed in the Western Hemisphere by American Elsevier, New York. viii+353 pp. $14.75. (Telegraphic Review, October 1969.)

This book is an excellent text for an introductory course in general topology for graduate students. The style of the exposition is well geared to the needs of such students, avoiding undue austerity and abstraction, and at the same time the numerous exercises provide the reader with both a guide to his understanding and practice in the manipulation of new concepts. In both respects, text and exercises, it is natural to compare this book with other standard treatments of the subject, say Bourbaki and Kelley. The present text (or at least its first 6 chapters) covers material comparable to that of Kelley, although convergence is treated, as in Bourbaki, by filters rather than nets. On the other hand the exercises in the text under review, while by no means trivial, are less demanding than those in either Bourbaki or Kelley. For example, in general, new concepts and additional theorems are not introduced here as exercises. The result is a text which is less encyclopaedic than Bourbaki or Kelley, but one which is possibly more practical for teaching beginners.

As regards contents, the material of the first six chapters follows a more or less standard pattern: I. Set theory, and motivation to topology by considering the topology of the plane. II. Topological spaces, open and closed sets, neighborhoods, convergence, mappings, product and quotient spaces, inverse limits and connectedness. III. Separation axioms, countability axioms, metrizability. IV. Compactness. V. Paracompact spaces. VI. Metrizability.

At this point the author, following (justifiably after all) his personal preferences, introduces some rather specialized advanced topics, including a section on P-spaces at the end of Chapter VI. The essential theorem here is that the product of a space R with every metric space is normal if and only if R is a normal P-space. Chapter VII discusses mapping spaces, with a long section on a very abstract theory of extension of mappings. The author points out himself that these special topics in Chapters VI and VII are not necessarily intended for a first reading.

There is an extensive bibliography of source material.

ANDREW WALLACE, University of Pennsylvania

Lectures on the Calculus of Variations and Optimal Control Theory. By L. C. Young. Saunders, Philadelphia, 1969. xi+331 pp. $15.00. (Telegraphic Review, October 1969.)

The appearance of this book is one of the most exciting events for friends of the Calculus of Variations since the publication of Carathéodory's classic in 1935 on the calculus of variations and partial differential equations of first order.

The author, who made a lasting impact on the development of this subject in 1933 when he conceived the notion of a generalized curve, gives here a very lively, greatly stimulating, and highly personalized account of the calculus of variations and optimal control theory. It is testimony to the far reaching thrust of his work that generalized curves resurfaced in recent years under the disguise of chattering controls in the modern theory of optimal control. The class of generalized curves contains, besides ordinary curves, other abstract objects that have an ordinary curve as trace but whose tangent line changes incessantly and with infinite rapidity between prescribed values. The initial motivation for the introduction of generalized curves was to provide solutions to variational problems that do not have solutions among the ordinary curves. The author credits Hilbert's statement that every problem of the calculus of variations has a solution, provided the word solution is suitably understood, with providing the proper motivation and stimulation for his own work. (Actually, Hilbert went even further in his essay *Naturerkennen und Logik*, where he stated that "The real reason for Comte's failure to find an unsolvable problem is, in my opinion, that an unsolvable problem does not, altogether, exist." This was, of course, written one year before Gödel's paper on undecidable propositions appeared.) It then developed that spaces of such generalized curves with a suitable topology provide the proper setting for the formulation and solution of the problem of existence of a solution to variational problems, and more recently, optimal control problems. The author and then E. J. McShane and others have demonstrated this for a variety of cases. In many instances, the solution turns out to be an ordinary curve after all.

This book is divided into two volumes, the first entitled *Lectures on the Calculus of Variations*, the other *Optimal Control Theory*. These, in turn, are subdivided into pre- and post-generalized curves segments. A list of chapter headings, though grossly inadequate for the purpose of describing the content of this book, will have to suffice here to convey some idea of its scope: The Method of Geodesic Coverings, Duality and Local Embedding, Embedding in the Large, Hamiltonians in the Large—Convexity—Inequalities and Functional Analysis, Existence Theory and its Consequences, Generalized Curves and Flows, The Nature of Control Problems, Naive Optimal Control Theory, The Application of Standard Variational Methods to Optimal Control, Generalized Optimal Control. There are two appendices to be found at the end of Volume I, one on Convexity and Integration and another on the Variational Significance and Structure of Generalized Flows. The only way to obtain accurate information and true appreciation for the content is to read the book.

The quest for existence theorems looms large in the foreground and not only because this was the principal motivating force for the author's distinguished research. It is, of course, futile to establish necessary and sufficient conditions which have to be satisfied by a solution of a problem unless one has an understanding of the nature of the solution and of its existence. To drive his point home, the author, on occasions, takes a very extreme position such as claiming repeatedly that necessary conditions (they are occasionally referred to as recipes) require no proof, or when he labels the various theories dealing with necessary conditions (embracing among others the theory of the first variation, the Lagrange multiplier rule, and the maximum principle of Pontryagin) as naive. Although this is not meant in a derogatory sense, it does seem rather harsh upon first reading, and this is precisely the author's intention. The shock wears off, however, as one becomes immersed in and persuaded by the author's modification and masterful justification of his position.

To the same extent as the author's approach to mathematics, is his approach to the teaching of mathematics highly original. The author pleads a strong case against the excesses in mathematical formalism and there is much prose to be found between formulas. At times, however, one might appreciate a less informal approach as, for example, in the definition of "fine convergence" which is basic to the notion of generalized curve, or when the author discusses the equivalence of implicit and parameter representation of varieties.

In his many refreshing asides, the author not only puts ideas and techniques into their historic perspective but also succeeds in making men, who for many of us are merely revered names, come alive through skillful selection of quotes and descriptions of their interaction with each other and the subject matter at hand. Sometimes his highly individual comments do not correspond to historical fact. For example, the name "Zorn's lemma" was not the result of Bourbakian caprice (p. 102), nor is it true that Cantor was confined to an asylum by his enemies (p. 103). The author also manages to convey to the reader, what some might find eccentric but what this reviewer finds most enjoyable, some of his very personal attitudes towards the real world, such as his distaste for boats that are propelled by an internal combustion engine, and his frustrations with the parlor game of musical chairs.

The fact that most ideas are first introduced by elaborate and well-illustrated discussions of simple problems might, upon superficial examination, create the impression that the book is easy to read. It is not. This is, of course, due to the sophisticated nature of a very difficult subject matter. Though the author states that these lectures were delivered to students of rather varied backgrounds and have been written for readers with only a minimum of mathematical preparation, but with an intense desire to learn, I would think that the reader should be well-versed in hard analysis and be prepared to learn basic and important facts in linear functional analysis which are developed and used mainly in the portions that deal with convexity and generalized curves. It should be noted that the author's terminology deviates somewhat from the

standard so that the reader with a background knowledge in functional analysis will have to familiarize himself with some new technical terms such as the "dutiful dual," the "B . . . i compact sets," the "Riesz mixture," and many others.

To summarize: A beautiful book—though not a textbook by any conventional standards—that is bound to stimulate many mathematicians and students of mathematics.

HANS SAGAN, North Carolina State University

C *Calculus—With an Introduction to Linear Algebra.* By John G. Hocking. Holt, Rinehart and Winston, New York, 1970. 866 pp. $14.75. (Telegraphic Review, January 1971.)

Our comments on Hocking's *Calculus* are based on only one semester's use. However, we feel that the book is of sufficient importance and offers enough special features that it is deserving of prompt commentary based on classroom experience. The book is being used for twelve sections of approximately thirty students each; three of these sections were designated "honors." Most of the students were from the top third of their high school classes and the honors students mostly have SAT mathematics aptitude scores of 700 or better. We covered six chapters in three class hours per week for thirteen weeks.

Hocking's *Calculus* has several notable features. The problem sets are divided between "exercises" and "problems" and there is a significant and, we feel, valuable difference. The exercises are mostly directly useful in mastering the basic material of the text and were appreciated in this capacity, especially by the nonhonors students. The problems involved another of the unusual aspects of this book; they were, where appropriate, divided into categories: mathematics, physical sciences, and economics. The mathematics problems were mostly difficult theory, too difficult for nonhonors students but challenging in a way which really interested the honors students (on a questionnaire they frequently made special note of this). The physics problems are typical and good with the exceptions noted below. The economics problems are in a class by themselves among calculus books. They are numerous and involve significant depth mathematically and economically to the point that an occasional paragraph defining concepts from economics is a necessity. In fact such paragraphs are provided as part of the problems.

The text starts the study of limits with a study of convergence of sequences, and then develops the usual $\epsilon-\delta$ theory quite fully. A chapter on series precedes integration. The differentiation, which comes before the series, and the integration are both developed without the aid of transcendental functions, which are introduced later. The linear algebra, though minimal at first, is truly integrated in the spirit of the text, but the real development comes later. This is really a three-semester book with multivariable calculus developing the concepts which are studied in the last chapter as formal linear algebra. The multivariable cal-

culus is mathematically weaker than the single variable calculus, but covers the fundamentals nicely for more advanced work.

The publisher has devoted a lot of time, energy, and expense (for the student) to format and color. But the students seem not to notice this and the display of definitions and other casual material for reference is often poor. The greatest fault is the unbelievable number of errors in the text; these are mostly obvious misprints. Also the answers to the problems are frustrating (a word used frequently on the questionnaires) to the student, because the answer given is much too frequently one obtainable by making a standard kind of mistake. Both must be remedied by a second printing (or edition).

The examples are numerous and usually good. They are somewhat unusual in the degree to which they develop theory rather than exemplify exercises, though none are typical of the hardest problems.

Further, less immediate difficulties occur in two matters of craftsmanship in the final editing. Hocking attempts to lead the students by means of the problems to future theory and applications, but this basically good idea is rendered nonproductive on a number of occasions by the fact that the student is completely overwhelmed or led to nearly useless and thoroughly messy *ad hoc* techniques. Secondly, a number of references to definitions or previous sections are found to be to nonexistent definitions or sections which have obviously been moved somewhere else in a later draft of the book.

On the whole, we and many of our students recommend the book as it is but would reserve a really strong recommendation until a second edition cures the above problems, most of which appear to be of a technical editorial nature.

LORRAINE L. KELLER and J. A. SEEBACH, JR., St. Olaf College

Simplified Independence Proofs. By J. Barkley Rosser. Academic Press, New York and London, 1969. xv+217 pp. $10.00. (Telegraphic Review, December 1969.)

In the early nineteen sixties Paul J. Cohen discovered a means of constructing a variety of different models of set theory and used these to prove, among other things, the independence of the continuum hypothesis and the axiom of choice. Shortly thereafter Robert M. Solovay showed that these proofs could be recast in the framework of a Boolean valued set theory, that is, a set theory in which statements, rather than being true or false, take on values in some prescribed Boolean algebra. It is this latter concept that the author deals with.

After an introductory chapter the author proceeds to define and discuss abstract topological spaces and Boolean algebras. The connection between them is then established with a proof that, under the proper operations, the family of regular open sets in a topological space forms a complete Boolean algebra. No prior knowledge on the reader's part is assumed and all proofs are written out in great detail. The author also uses this chapter to introduce the particular spaces and their associated algebras, which he will use later, as well as the notions of

automorphisms of Boolean algebras, groups of such automorphisms, and what he refers to as filters of groups of automorphisms.

In Chapter 3 the basic transfinite inductive procedure for constructing a Boolean valued model of set theory from a given Boolean algebra and a given filter of groups of automorphisms on that algebra is presented. (In the preface this model is said to resemble one described in some notes by Kenneth Kunen.) At each level the objects added are those functions from the previously added functions (the first level consists of the empty function) into the algebra which satisfy certain internal consistency requirements and certain requirements concerning the automorphisms. Then for any two elements a and b of the model the statements "$a\in b$" and "$a=b$" are assigned values from the algebra. This assignment of Boolean values is then extended in a canonical way to all statements of the appropriate language. Much of the remainder of the chapter consists of proofs that all the axioms of the first order predicate calculus and all of the axioms of Zermelo-Fraenkel set theory (including even the axiom of choice if the filter is simple enough) have been assigned value 1. But since it is also shown that valuation 1 is preserved under *modus ponens* (i.e., if A and $A\rightarrow B$ are statements to which the value 1 has been assigned then B will also have been assigned value 1) it follows immediately that every theorem of Zermelo-Fraenkel will have been assigned value 1. Thus to prove that a given statement is not a theorem, it is sufficient to find a particular Boolean algebra and a particular filter on the algebra such that with respect to the associated model the statement is assigned a value other than 1. Also discussed in this chapter is a means of extending certain automorphisms of the algebra to automorphisms of the model and an embedding of what might be called the real or intuitive sets into the model. This latter then leads to a discussion of cardinals and ordinals in the model.

Most of the remaining chapters are devoted to the study of particular models and, more specifically, to the properties of the set of "real numbers" in these models. In Chapter 4 an algebra and filter are presented such that in the associated model every formula containing only two free variables and no constants fails to well order the reals (i.e., if $F(x, y)$ is any such formula, then the statement which says that F well orders the reals does not have value 1). Thus the axiom of constructibility is shown to be not a theorem and therefore, by Godel's work, independent. Similarly in Chapter 6 a model is presented in which the reals cannot be well ordered at all thereby proving that the axiom of choice is independent. These proofs both depend upon a careful study of the properties of those elements of the model which are invariant under automorphisms connected with the defining filter.

In Chapters 7, 8, and 9 various forms of the negation of the generalized continuum hypothesis are proven to be consistent; here the proofs are combinatorial in nature and the filters are not used at all.

Finally, there is a chapter (5) covering the relationship between Boolean

valued set theory and forcing, and a concluding chapter dealing with some technical and conceptual problems.

Overall the reviewer found the book most interesting and impressive. There is a great mass of material (in Chapters 2 and 3 alone there are 88 theorems) to be checked and almost all of the necessary proofs are carefully and completely worked out. While the casual reader may wish to skip over many of these, there is certainly a real need to have them appear written out somewhere in the literature, a need this book admirably fulfills.

In a great many cases the more difficult results are obtained by using essentially the same techniques as are used in proving the analogous results concerning models constructed using forcing. This is both an advantage and a disadvantage. The reader familiar with standard forcing arguments will have no trouble making the transition and will be easily able to supply the necessary motivation. On the other hand, the reader not familiar with such arguments may frequently find himself confronted with proofs which can be checked step by step but which nevertheless seem completely arbitrary. Thus while the set theorist will undoubtedly be interested in most if not all of the book, the casual reader would do well to follow the advice given by the author in the first chapter and attempt only certain sections. However, the author is perhaps overly optimistic; the reader should not be surprised to find himself skipping over even more detail than is recommended.

It should be mentioned that one definite prerequisite for reading this book is a complete familiarity with the first order predicate calculus. It is continually necessary to express moderately sophisticated concepts such as equinumerosity and cofinality directly in this language and the resulting statements are, even to one familiar with the procedure, difficult to decipher. Furthermore, the author uses a dot notation which is not explained in the book; the reader is referred to his earlier volume *Logic for Mathematicians*.

To sum up, Professor Rosser is to be highly commended for taking a new concept and making it, if not easily accessible, then at least as accessible as is perhaps possible to a general audience.

S. H. HECHLER, Case Western Reserve University

Elêments d'Histoire de Mathêmatiques. By Nicolas Bourbaki. Second edition, revised, corrected, and extended. Hermann, Paris, 1969. 323 pp. 36F. (Telegraphic Review, November 1969.)

This second edition differs from the first one (1960) in the addition of two topics, "Commutative Algebra. The Theory of Algebraic Numbers" (29 pp.) and "Haar Measure and Convolutions" (6 pp.), the expansion of the bibliography to accommodate the references on these two topics, and the inclusion of an index of names. The original 21 topics are unchanged.

For the benefit of those readers who may have missed the review of the first edition we might note that this history consists of the historical sections in the

numerous Bourbaki texts that have thus far appeared. Consequently the articles are unrelated to each other and cover only some topics in the history of mathematics. The articles are very tightly written and the accounts are framed in modern concepts and terminology. This last fact is unfortunate because it creates a false impression of how the earlier mathematicians thought and worked. Like most historical articles and texts this one does not explain the concepts and theorems that are treated therein. One must know the mathematics involved to appreciate the history. Because the articles are disconnected and are limited to those subjects already covered in the Bourbaki series, they cannot offer any broad picture of the development of mathematics. They are rather pieces of a jigsaw puzzle which the reader must fit into the broad picture.

Despite these limitations the book is valuable because it is the only one that gives any account at all of modern developments. Some of the articles are extensive enough to enable a specialist in a particular field to acquire the essential historical background for his area of research. The references to the original literature are very helpful.

MORRIS KLINE, New York University

Greek Mathematical Thought and the Origin of Algebra. By Jacob Klein. Translated by Eva Brann. M.I.T. Press, Cambridge, Mass., 1968. xv+360 pp. $12.50. (Telegraphic Review, June/July, 1969.)

In an admirable attempt to avoid the pitfalls of those writers who have "looked at the problems of the origins of modern algebra from the modern period back," Klein begins by examining various Greek interpretations of the meaning of "arithmetic," "logistic," "number" and related concepts. Especially important is his study of Diophantus's *Arithmetic*, which Klein interprets, perhaps incorrectly, as a purely Greek product. Skipping from Diophantus to the late sixteenth century, Klein shows how Viète and later mathematicians misinterpreted and reinterpreted Diophantus until at last, with Wallis, "the whole complex of ontological problems which surrounds the ancient concept of number loses its object in the context of the symbolic conception," and the modern concept of number appears.

This is not a book for readers with only a casual interest in the subject, since ideas which in any event are not simple are made more difficult by a confusing format and the inclusion of much Greek and Latin. Klein might also be faulted for concentrating on only a few mathematicians and giving insufficient credit to those Arabic and European authors who helped to make possible the achievements of Viète and his successors. However, for readers who are willing to overcome or ignore these flaws, Klein's book presents extremely valuable insights into some Greek and early modern views of the concept of number. It is also one of the few sources that provides a thoughtful introduction to Viète's mathematics.

BETTY R. ESTES, Fairleigh Dickinson University

TELEGRAPHIC REVIEWS

Telegraphic reviews are intended to give prompt notice of books, with sufficient information to assist in deciding whether to order an examination copy or suggest library purchase. Possible uses are indicated as follows:

B = college bookstore stock L = library purchase
P = professional reading S = supplementary reading
T = textbook E = teacher education
13 to 18 = freshman to second year graduate level usage
1 to 4 = approximate time in semesters to cover text
* = positive emphasis ? = negative emphasis

Books on high school material (pre-calculus) are denoted REMEDIAL, and normally receive telegraphic reviews only if they are written for college students. Publishers are denoted by the standard abbreviations used in *Books in Print*, which gives complete addresses.

ADVANCED CALCULUS, T(15-16: 1, 2), L. *Advanced Calculus. Functions of Several Variables. Monografie Matematyczne, Tom 51.* Roman Sikorski. Krieger, 1969, 460 pp, $16.50. At the same level as Goffman and Fleming, this work also appeared in 1965 (in Polish). Much more is done with the theory of integrals with respect to Lebesgue measure than in Goffman or Fleming. In particular, integrals on hypersurfaces and polyhedra are treated. The necessary ideas from set-theory, algebra, and topology are introduced in chapters prefatory to the body of the text. W.C.R.

ALGEBRA AND FUNCTIONAL ANALYSIS. *Ten Papers on Algebra and Functional Analysis. American Mathematical Society Translations, Series 2, Volume 96.* AMS, 1970, 254 pp, $13.20. Papers by Ševrin, Filippov, Rozen, Vinogradov, Šaĭn, Garkavi, and Ginzburg. R.J.

ALGEBRA, LATTICES, T(17), P, L. *Theory of Symmetric Lattices.* P. Maeda and S. Maeda. *Die Grundlehren der mathematischen Wissenschaften, Band 173.* Springer-Verlag, 1970, 190 pp, $13.20. A research monograph on modular and M-symmetric lattices. For both the general and atomistic cases there is a theory of geometric symmetric lattices and a theory of analytic symmetric lattices. There are problems included and the book could be used as a seminar text. W.C.R.

ALGEBRA, LINEAR, T*(14: 1), S, L. *Linear Algebra, Volume 2A.* Michael O'Nan. Harcourt Brace Jovanovich, Inc., 1971, 385 pp, $9.95. This volume in the Eagle Mathematics Series gives a thorough and lucid presentation of elementary linear algebra which proceeds from concrete topics (linear equations, column vectors, matrices and determinants) to more abstract topics (vector spaces, linear transformations, inner products, eigenvalues, and canonical forms). Throughout the presentation, concepts and examples are related to their geometric counterparts while care is taken to provide detailed proofs of the theorems. The numerous exercises range from routine computation to challenging theoretical problems. J.N.C.

ALGEBRAIC GEOMETRY, T*(17: 1), S, L. *Multidimensional Analytic Geometry. Monografie Matematyczne, Tom 50.* Karol Borsuk. Krieger, 1969, 443 pp, $16.50. A rigorous, completely self-contained

development of finite dimensional Cartesian and projective spaces from linear algebra. There is no elementary geometry used. The Erlangen Program is followed. There are few problems and they are of a straightforward nature. W.C.R.

ALGEBRAIC TOPOLOGY, T(17-18: 2), *Algebraic Topology, Homology and Cohomology*. Andrew H. Wallace. Benjamin, 1970, 272 pp, $12.95. A classical (without category theory) introduction to homology and cohomology theory. The orientation is refreshingly geometric though the (algebraic) theory of chain complexes is included. Singular and Čech theories are given priority. No duality theory, cohomology operations (other than cup product), or anything beyond the basics. J.A.S.

ANALYSIS, P, L*, *Twelve Papers on Real and Complex Function Theory. American Mathematical Society Translations, Series 2, Volume 88.* AMS, 1970, 325 pp, $16.50. Contains three papers on convex functions, two on univalent functions, one on differentiable functions in n-space, and six on entire functions. More than half the book is devoted to four papers by A.A. Gol'dberg on "An integral with respect to a semiadditive measure and its application to the theory of entire functions." T.A.V.

ANALYSIS AND QUANTUM MECHANICS, P, L, (RESEARCH), *Eighteen Papers on Analysis and Quantum Mechanics. American Mathematical Society Translations, Series 2, Volume 91.* AMS, 1970, 310 pp, $15.80. R.J.

ATOMIC PHYSICS, S, P, L*, *Problems of Atomic Dynamics*. Max Born. MIT Pr, 1970, 200 pp, $2.95 (P). A reprint of a 1926 lecture series, the book is of interest to physicists. Contents are dated but the style, notation, and level are excellent for today's students of atomic physics. The book gives an excellent transition from the "old" quantum to the "new." B.C.

CALCULUS, T(13: 2), *Introduction to Calculus 1 and 2*. Alfred B. Willcox, R. Creighton Buck, Henry G. Jacob, and Duane W. Bailey. Houghton Mifflin, 1971, 669 pp, $12.95; *Solutions Manual for Introduction to Calculus 1.*, 123 pp, $1.50 (P); *Solutions Manual for Introduction to Calculus 2.*, 112 pp, $1.25 (P). A genuine attempt at presenting the calculus with a fresh approach. Problems and examples precede many discussions to motivate and amplify ideas. Probably the feature of most interest is the presentation of material in Part I, then returning to the ideas to sharpen techniques and extend concepts in the second semester; Part II has chapter headings such as "Limits Revisited," "Derivatives Revisited," and "Integrals Revisited." L.L.K.

CALCULUS, B, L, *A New Table of Indefinite Integrals, Computer Processed*. Melvin Klerer and Fred Grossman. Dover, 1971, 198 pp, $3 (P). Just what every student of integral calculus has been asking for: over 2,000 well organized and apparently extremely accurate integration formulas. They were manually collected from previous tables; then computers were used to symbolically differentiate them and numerically check the results, and were also used for type-setting. R.W.N.

CALCULUS, T*(13: 1), *A Short Course in Calculus, 2nd ed.* Jack G. Ceder and David L. Outcalt. Worth, 1971, 341 pp, $8.95. Selected

topics from the calculus for students of biology, business, economics, psychology, and sociology. Included are sequences, max-min problems for functions of one or more variables, improper integrals, and some differential equations. There are more exercises and examples than the first edition, and many topics have been moved to the appendix with additional material for reference or review also in the appendix. L.L.K.

CALCULUS AND LINEAR ALGEBRA, T(14: 2). *Calculus and Linear Algebra. Vector Spaces, Many-Variable Calculus, and Differential Equations, Volume II.* Wilfred Kaplan and Donald J. Lewis. Wiley, 1971, 581 pp, $10.95. See telegraphic review of first volume, published in Volume 77, No. 7. The second volume competes with such books as the recent ones by Williamson, Crowell and Trotter, and by Osserman. Like the former, it treats linear algebra first and then uses it extensively, but it will perhaps be easier for students to read. Like the latter, it proves most calculus theorems only for functions of two variables, but it gives proofs of some major theorems Osserman merely states. It would be hard going to cover the book, even without the chapter on differential equations, in two semesters. Deserves the serious consideration of all those who believe linear algebra and multivariate calculus should be integrated but are not satisfied by any of the earlier efforts to do so in an elementary text. J.D.-B.

COMBINATORICS, MATROIDS, T(16-17: 1), P, L. *Introduction to the Theory of Matroids: Modern Analytic and Computational Methods in Science and Mathematics, Number 37.* W.T. Tutte. Am Elsevier, 1971, 84 pp, $7.50. This monograph offers a direct and clear presentation of the basic theory of matroids. It does not include applications, examples, or exercises. L.C.L.

COMPLEX ANALYSIS, T***(15-17: 1, 2), B, L. *Complex Variables.* Norman Levinson and Raymond M. Redheffer. Holden-Day, 1970, 429 pp, $13. Intended for a one or two semester course at the senior or early graduate level, this book appears to be one of the very best texts to appear in this area. The development of the theory is concise at a level of generality suitable for applications. Each chapter gives a brief logical development of the theory and has a section on the applications thereof. Very well written with an abundance of examples and problems of varying difficulty, this book may well become the standard text in complex analysis at this level. T.A.V.

COMPLEX VARIABLE, T(16-17: 1), L. *A First Course on Complex Functions.* G.J.O. Jameson. B & N, 1970, 148 pp, $5 (P). "This book contains a rigorous coverage of those topics (and only those topics) that, in the author's judgment, are suitable for inclusion in a first course on Complex Functions. Roughly speaking, these can be summarized as being the things that can be done with Cauchy's integral formula and the residue theorem." The author seems to have succeeded in presenting the theory in a precise, elegant, and appreciative way. In paperback. R.B.K.

COMPUTER SCIENCE, T?, L?. *Modern Programming: Fortran IV.* Henry Mullish. Ginn-Blaisdell, 1968, 132 pp, $5.25 (P). Another Fortran manual. Not modern (7094 Fortran IV). Weak on subprograms, no bibliography, few problems. J.G.L.

COMPUTER SCIENCE, T***, L*, B***, *Rudiments of Fortran.* Loren P. Meissner. A-W, 1971, 109 pp. Enough Fortran for a beginner, without fancy formatting, etc. A primer, not a complete language manual. No mathematical background assumed, so is weak on scientific problems. Only SQRT, ABS, IABS, INT, FLOAT and MOD as standard functions, the last three well explained and often used. With supplementary problems, would be useful where a full programming course is not required or available. J.G.L.

COMPUTER SCIENCE, T, L, B, *Basic for Beginners.* Wilson Y. Gateley and Gary G. Bitter. McGraw, 1970, 152 pp, $3.95 (P). Self-teaching manual on slightly extended BASIC. Similar in amount of material on syntax to original Kemeny and Kurtz. More detailed on use of an ASR 33, etc. (details which change so drastically that they must be supplied locally anyway), fewer and less interesting problems and complete programs than Kemeny and Kurtz. Includes "ON ... GO TO", string I/O, omits "GOSUB" and "RETURN." J.G.L.

COMPUTER SCIENCE, T, L*, B, *APL/360: An Interactive Approach.* Leonard Gilman and Allen J. Rose. Wiley, 1970, 335 pp, $6.95 (P). From a series of videotape lectures for IBM employees. Designed to be used with immediate access to an IBM 360 supporting APL and not much use without it. Weak on problems, strong on the virtues of APL. Structure requires many leaps of faith. J.G.L.

COMPUTER SCIENCE, L, E, B, *Computing and Computer Science: A First Course With Fortran IV.* Macmillan, 1970, 398 pp. *Computing and Computer Science: A First Course with PL/I.* T.D. Sterling and S.V. Pollack. Macmillan, 1970, 414 pp. Identical texts. Covering many elementary technical aspects of computing plus enough of the syntax of PL/I or Fortran IV to give the student some mastery of programming. More technical, less algorithmically focused than several similar multi-language texts. See TR June, August 1970. J.G.L.

COMPUTER SCIENCE, T*, L, *Computer Science. Basic Language Programming.* Alexandra I. Forsythe, Thomas A. Keenan, Elliott I. Organick and Warren Stenberg. Wiley, 1970, 124 pp, $3.95 (P). Basic language supplement to the same authors' *Computer Science: A First Course* or *A Primer* (see extended review, January 1971). Designed to mirror central text, not self-contained, similar Fortran, PL/I and APL supplements available. J.G.L.

DICTIONARY, L?, *A Dictionary of Named Effects and Laws in Chemistry, Physics, and Mathematics.* D.W.G. Ballentyne and D.R. Lovett. Chapman and Hall, 1970, 335 pp, $9.50. This dictionary contains explanations of what certain physical, chemical, and mathematical entities are, namely those that use the name of a person to label them. The mathematical explanations are frequently misleading. The work does represent a source where one can obtain some idea of what certain "named" concepts are. R.J.

DIFFERENTIAL AND INTEGRAL EQUATIONS, T(17-18: 2), L, *Equations of Mathematical Physics.* V.S. Vladimirov. Ed: Alan Jeffrey. Transl: Audrey Littlewood. Marcel Dekker, 1971, 418 pp, $19.75. A non-traditional, fresh look at the classical boundary value problems for differential equations of mathematical physics. The approach is to use ideas from distribution theory: a "generalized solution" is defined with use of the concepts of generalized functions and

derivatives. After having done this, the book considers specifically the generalized Cauchy problem for the wave and heat conduction equations, boundary value problems for elliptic equations, and the mixed problem for hyperbolic and parabolic equations. A chapter on the theory of integral equations with a polar kernel is included. The book is a welcome translation. D.F.A.

DIFFERENTIAL EQUATIONS, P, L, *Boundary Problems for Differential Equations II. Proceedings of the Steklov Institute of Mathematics, Number 103.* Ed: V.P. Mihailov. AMS, 1970, 213 pp, $18.60 (P). This volume consists of eleven papers given in a mathematical physics seminar at the Steklov Institute of Mathematics, U.S.S.R. Nearly half the papers deal with the stabilization of solutions of nonstationary boundary value problems as $t \to \infty$. The other papers are devoted to some boundary value problems for elliptic equations. T.A.V.

DIFFERENTIAL EQUATIONS AND LINEAR ALGEBRA, T(14: 1), *Linear Mathematics.* Philip Gillett. Prindle, 1970, 373 pp, $9.95. A book that brings together linear algebra and differential equations at the sophomore level without overwhelming the student. The style is unusual and the author has an irresistible sense of humor. His claim: "the level of abstraction is somewhere between that of a good calculus course and Herstein's *Topics in Algebra*; thus the climb is Alpine but not Himalayan." L.L.K.

ECONOMICS, T(15-16), L, *Economic Growth and Development: A Mathematical Introduction.* Philip A. Neher. Wiley, 1971, 322 pp, $9.95. A text in economic growth and development using two basic single-sector models: a neoclassical model of an advanced economy and a dynamic model of a primitive economy. The mathematics is extremely intuitive; the text includes instruction in the rules of differential calculus and elementary differential equations. The mathematician may profit from the intuitive discussions, learning more of the application of elementary calculus to economics and the analysis of mathematical models. However, the book is directed to a mathematically unsophisticated reader. R.B.K.

EDUCATION, T*(14-16: 2), E(ELEMENTARY), *Principles of Arithmetic and Geometry for Elementary School Teachers.* Carl B. Allendoerfer. Macmillan, 1971, 672 pp, $9.95. Structure of the number system and informal geometry. Arithmetic chapters focus on presenting ideas behind mechanics. Informal geometry includes geometric transformations. Each chapter preceded by a Readiness Test and followed by programmed exercises. Summary Tests for each of five parts. Thirteen supplementary films available. The book is based on class tested revisions of CEM initiated "multimedia" presentations of a Level I course. Instructor's Manual. P.J.

EDUCATION, ELEMENTARY, E(2), *Basic Concepts of Elementary Mathematics.* John M. Peterson. Prindle, 1971, 435 pp, $9.95. Covers the topics recommended by CUPM for the first two courses for prospective elementary school teachers, plus chapters on geometry and probability. Each chapter that introduces a number system is followed by a chapter that applies the properties of that number system. J.N.C.

FINITE MATHEMATICS, T(13: 2). *Finite Mathematics with Applications.* A.W. Goodman and J.S. Ratti. Macmillan, 1971, 490 pp, $10.95. A text designed to present a slice of mathematics that is interesting, meaningful and useful without involving the calculus. Included are the usual topics: theory, logic, sets, combinatorial analysis, probability, vectors and matrices, with applications ranging from game theory to graph theory. L.L.K.

FOUNDATIONS, LOGIC, T(13: 1). *Sets and Logic.* Samuel C. Hanna and John C. Saber. Richard Irwin, 1971, 274 pp, $4.95. A presentation of the fundamentals of sets and logic that could easily be taught as a first course in college mathematics. It is written in a style which a non-mathematician could handle, and includes over 100 examples and 250 exercises. L.L.K.

FUNCTIONAL ANALYSIS, P, L(RESEARCH). *Nine Papers on Functional Analysis. American Mathematical Society, Series 2, Volume 93.* AMS, 1970, 253 pp, $13. R.J.

FUNCTIONS OF REAL VARIABLES, T(18: 1, 2), P, L. *Singular Integrals and Differentiability Properties of Functions.* Elias M. Stein. Princeton U Pr, 1970, 287 pp, $11. An examination of the unity which exists in these areas of analysis: the existence and boundedness of singular integral operators, the boundary behavior of harmonic functions, and the differentiability properties of functions of several variables. Topics include covering lemmas, maximal functions, the Marcinkiewicz interpolation theorem, singular integrals generalizing the Hilbert transform, harmonic functions represented as Poisson integrals, Littlewood-Paley theory, multipliers, Sobolov spaces and their variants, extension theorems, conjugate harmonic functions, and almost-everywhere differentiability theorems. Prerequisite: elementary integration and Fourier transform theory. D.F.A.

*GENERAL, T(13-16), S, L. *Ingenuity in Mathematics.* Ross Honsberger. Random House, 1970, 204 pp. This is #23 of the New Mathematical Library Series for The Monograph Project of SMSG. A well written set of 19 essays on topics from number theory, geometry, combinatorics, logic, and probability. Exhibits elegant and ingenious approaches to mathematical thinking. Modelled after Rademacher and Toeplitz's *The Enjoyment of Mathematics.* A.G.

??GENERAL, DICTIONARY. *Dictionary of Mathematics.* T. Alaric Millington and William Millington. B & N, 1966, 2nd ed, 1971, 259 pp, $2 (P). "Mathematics" herein includes business arithmetic, high school physics, and surveying as well as high-school mathematics (with a British accent). The scattering of entries from calculus, algebra, topology and logic includes so many errors that a knowledgeable freshman would tear his hair (you might look up infinitesimal; Rolle's theorem; homomorphism; cardinal number). As the authors say, "The need for clear thinking and clarity of expression has never been greater." L.A.S.

HISTORY, S, P, L. *Geschichte der Mathematik.* A.G. Kästner. Georg Olms Verlag, 1970. *Volume I*, 708 pp; *Volume II*, 759 pp; *Volume III*, 484 pp; *Volume IV*, 556 pp, $79.80. A facsimile of the original (1796-1800) plus an informative note and name index by J.E. Hofmann, the dean of West German historians of mathematics. In spite of

Kästner's weaknesses as a mathematician (his most distinguished student, Gauss, lampooned him as the leading poet among the mathematicians and the leading mathematician among the poets) and as a historian, the work is a valuable source of first-hand descriptions of many rare mathematical works. K.O.M.

HISTORY, PHILOSOPHY, S, P, L, *The Usefulness of Mathematical Learning*. Isaac Barrow. Frank Cass, 1970, 458 pp, £7.7s. A facsimile reprint of the 1734 translation from Latin of his inaugural oration as Lucasian professor and of twenty-three lectures delivered to students at Cambridge during 1664-1666 on a variety of metamathematical topics, e.g. the nature and subdivisions of mathematics, the "identity of arithmetic and geometry," and proportions versus numbers. An important source for historians and for the pleasure of any mathematician with leisure to browse. There is a portrait and both topical and subject indexes. K.O.M.

LINEAR ALGEBRA, T(13: 1), S, *Basic Linear Algebra*. B.C. Tetra. Har-Row, 1971, 136 pp, $2.95 (P). This is a very elementary approach to Linear Algebra for the first year college student. It is an inexpensive paperback, which is a feature that makes it adaptable to a variety of uses. It has three main chapters: Matrices and Vectors, Applications, and Linear Spaces. L.L.K.

NONLINEAR PROGRAMMING, S, P*, L, *Nonlinear Programming*. Ed: J.B. Rosen, O.L. Mangasarian, and K. Ritter. *Proceedings of a Symposium Conducted by the Mathematics Research Center, The University of Wisconsin, Madison, 1970.* Acad Pr, 1970, 490 pp, $10.50. Seventeen papers which "emphasize those algorithms and related theory which lead to efficient computational methods for solving nonlinear programming problems." The typescript printing is justified by the expeditious publishing of these current papers. R.W.N.

NUMERICAL METHODS, ORDINARY DIFFERENTIAL EQUATIONS, T(15-16: 1), S, P, L, *Numerical Solution of Ordinary Differential Equations*. Leon Lapidus and John H. Seinfeld. Acad Pr, 1971, 299 pp, $16.50. A practical aid in the selection from among the many single-step, multi-step, and predictor-corrector methods for the numerical solution of initial value problems. No exercises are given explicitly; however,many are suggested by the advice-giving sections on numerical experiments and on published numerical results. Included are a chapter on the role of stability and short chapters on extrapolation methods and on attempts to adapt the ordinary methods to stiff equations. R.W.N.

ORDINARY DIFFERENTIAL EQUATIONS, T*(14-15: 1), *Topics in Differential Equations*. Allen D. Ziebur. Dickenson, 1970, 307 pp, $9.95. What is unusual about this text is not what topics are presented (these include: scalar first order initial value problems, scalar linear equations of the second order, linear systems of equations, phase plane analysis and stability, and an introduction to boundary value problems and Fourier series), but that coursing through it is an introductory course in using the computer to attack problems and understand theory. Most sections provide some theory, some examples, and a computer program for still another example; while a "computer-free" course using this text can be constructed, doing this would seem a shame--the text as is would be the basis for an exciting course. Its prerequisite: the standard

calculus sequence; no familiarity with the computer is assumed. D.F.A.

PHYSICS, P, L. *Lectures on Elementary Particles and Quantum Field Theory: 1970 Brandeis University Summer Institute in Theoretical Physics, Volume I.* Ed: Stanley Deser, Marc Grisaru, and Hugh Pendleton. MIT Pr, 1970, 592 pp, $16.95. Typescript prepared for quick availability to the physics community. Probably not mathematical enough to be of interest to any but a few mathematicians who want some contact with high level modern theoretical physics. J.A.S.

PROBABILITY, S**, P, L. *Probability and Related Topics in Physical Sciences. Lectures in Applied Mathematics, Proceedings of the Summer Seminar, Boulder, Colorado, 1957, Volume I.* Mark Kac. AMS, 1959, 266 pp, $10.10. "An expanded version of twelve lectures delivered at the Seminar in Applied Mathematics held in Boulder, Colorado in the summer of 1957." A fascinating exposition on probabilistic reasoning and techniques, probability in classical statistical mechanics, and integration in function spaces. Appendices include lectures by G.E. Uhlenbeck, A.R. Hibbs, and Balth. van der Pol. F.L.W.

PROBABILITY, T**(15-16: 1, 2), S. *Probability Theory and Applications.* Meyer Dwass. Benjamin, 1970, 413 pp, $12.95. For a post-calculus introduction to the field. An elementary treatment of stochastic processes is included. The book consists of the first 11 chapters of the author's *Probability and Statistics*. F.L.W.

PROBABILITY AND STATISTICS, T*(13: 1, 2), E, S. *Introduction to Statistics: A Fresh Approach.* Gottfried E. Noether. Houghton Mifflin, 1971, 253 pp, $9.95. Attempts to emphasize basic statistical ideas by using non-parametric methods before discussing more standard procedures. Presupposes only high school mathematics. F.L.W.

PROBABILITY AND STATISTICS, T(15-17: 1, 2). *Probability and Random Processes: An Introduction for Applied Scientists and Engineers.* Wilbur B. Davenport, Jr. McGraw, 1970, 542 pp, $14.95. "Directed towards students mainly interested in applications"; but presents "the underlying mathematical issues in a readable and technically honest way." The problems given are <u>all</u> to be worked by the student. Answers are available in a separate booklet. Presupposes calculus for Chapters 1-8 and Fourier analysis for the rest. F.L.W.

PROBABILITY AND STATISTICS, T**(15-17: 2, 3), S. *Probability and Statistics.* Meyer Dwass. Benjamin, 1970, 635 pp. The usual topics for a post-calculus course, plus unusual coverage of combinatoric problems and stochastic processes. Contains a chapter on linear algebra which is used in dealing with linear models. F.L.W.

REAL ANALYSIS, ELEMENTARY, T(13: 1). *Introduction to Abstract Mathematics.* T.A. Bick. Acad Pr, 1971, 217 pp, $8.50. One answer to the perennial question "when do they first meet a proof" is here provided by a thorough study of the real numbers (Peano to Dedekind) intended for the end of the freshman year. This is smoothly written and ought to be accessible to average-or-better students. Less commonly, the author concludes with a brief introduction to metric spaces, sufficient for him to discuss Moore-Smith convergence and

show how the properties of the reals have influenced the calculus courses. L.A.S.

REAL ANALYSIS, CALCULUS, T(13: 2), *The Calculus: An Introduction.* Casper Goffman. Har-Row, 1971, 422 pp, $9.95. The author claims to present calculus with "complete justification but without formal treatment." It is brief and exercises are presented with less imagination than Protter and Morrey (if that is possible). L.L.K.

REFERENCE, P, L**, B, *Formulas and Theorems in Pure Mathematics.* G.S. Carr. Chelsea, 1970, 971 pp, $12.50. A reprint of one of the greatest collections of mathematical results ever published with introduction by Jacques Dutka and with "a slight change of notation". There are 6165 formulas and theorems, well cross referenced. The detailed index incorporates a topical index of the contents of the 32 leading mathematical journals from 1800 to 1885! Generations of Cambridge wranglers used this book and Ramanujan based his work on it. K.O.M.

REMEDIAL, T(13: 1), *Introductory Algebra for College Students.* Eugene Nichols. HR & W, 1971, 463 pp, $9.95. Written for college students who have not had an algebra course in high school, it contains topics necessary for studying college algebra. J.N.C.

REMEDIAL, T(12: 2), *Algebra and Trigonometry.* Gordon Fuller. McGraw, 1971, 543 pp, $10.50. Another pre-calculus text with nothing to distinguish it from the hordes of others. L.L.K.

REMEDIAL, T(13: 1), *College Algebra and Trigonometry.* Margaret F. Willerding and Stephen Hoffman. Wiley, 1971, 500 pp, $9.95. A standard treatment of the usual topics plus determinants; sequences and series; permutations, combinations and probability. Omits function composition and algebraic operations on functions and could use graphical illustrations of a greater variety of functions. J.N.C.

REMEDIAL, T(1), *Number Systems: An Intuitive Approach.* Rex L. Hutton. Intext Educ, 1971, 361 pp, $8.95. Intuitive treatment of the number system from beginning set theory to the rationals, with a nod toward completeness of the reals. Not inspired. A.G.

REMEDIAL, T(13: 1, 2), S, *Intermediate Algebra for College Students.* Mary P. Dolciani, Robert H. Sorgenfrey and Edwin F. Beckenbach. Houghton Mifflin, 1971, 428 pp, $8.95. Remedial algebra. What more can be said? It contains the material normally found in a beginning algebra course in high school. The final chapters get to some topics which could be termed intermediate: Permutations, Combinations, and Probability; Exponential and Logarithmic Functions; and Matrices and Determinants. L.L.K.

SHEAF THEORY, T(18: 1, 2), P, L, *Garbentheorie.* R. Kultze. B.G. Teubner Stuttgart, 1970, 179 pp. A standard but fairly complete introduction to sheaf theory, assuming some knowledge of point-set topology and the theory of functions of one complex variable, but none of homological algebra. After an extensive development of cohomology the book ends with some elementary applications of Čech cohomology to the theory of functions of several complex variables. Problems at the end of each chapter, and a bibliography. J.D.-B.

STATISTICS, T(16-17: 2, 3), S, L. *Statistical Design and Analysis of Experiments.* Peter W.M. John. Macmillan, 1971, 356 pp, $14.95. Latin square, 2^n and 3^n factorial, incomplete block, and partially balanced designs. Fractional factorials and response surfaces. Presupposes linear algebra and a first course in mathematical statistics. Uses no measure theory. Extensive bibliography. F.L.W.

STATISTICS, T(16-17: 1, 2), S, P, L. *Sequential Tests of Statistical Hypotheses.* B.K. Ghosh. A-W, 1970, 454 pp, $15. Sequential probability ratio tests. Sequential tests for multi-parameter families, analysis of variances, and nonparametric tests. Background chapters on probability, classical hypothesis testing, and stochastic processes. Use of measure theory restricted to the appendices. F.L.W.

STATISTICS, DISTRIBUTION THEORY, P*, L*. *Distributions in Statistics.* Norman L. Johnson and Samuel Kotz. 3 volumes, Houghton Mifflin: *Discrete Distributions*, 1969, 328 pp, $12.95; *Continuous Univariate Distributions-1*, 1970, 300 pp, $13.95; *Continuous Univariate Distributions-2*, 1970, 306 pp, $13.95. These three volumes contain an impressive collection of facts about the distributions that occur most often in statistics (within the categories designated). Included, when appropriate, is information on historical development, moments and other properties, characterizations, approximations, related distributions and estimation of parameters (but nothing on hypothesis testing). The treatment is detailed, but concise and technical. Proofs and derivations are for the most part omitted but referenced. Also referenced are some of the applications that have been made of these distributions and sources of existing tables. This is a very thorough, well-documented work. R.S.K.

STOCHASTIC PROGRAMMING, S, P*, L. *Lecture Notes in Operations Research and Mathematical Systems-23: Foundations of Non-stationary Dynamic Programming with Discrete Time Parameter.* K. Hinderer. Springer-Verlag, 1970, 160 pp, $4.40 (P). Based on a summer course at the University of Hamburg. Gives a rigorous foundation to stochastic dynamic programming using $\bar{p}$-optimality in countable and general state spaces. R.W.N.

TOPOLOGY AND DIFFERENTIAL GEOMETRY, P, L. *Seventeen Papers on Topology and Differential Geometry. American Mathematical Society Translations, Series 2, Volume 92.* AMS, 1970, 284 pp, $14.40. R.J.

Reviewers Whose Initials Appear Above

David F. Appleyard, Carleton; Judith N. Cederberg, St. Olaf; Barry Cosens, St. Olaf; John Dyer-Bennet, Carleton; Arthur Gropen, Carleton; Richard Jarvinen, Carleton; Paul Jorgensen, Carleton; Lorraine L. Keller, St. Olaf; Roger B. Kirchner, Carleton; Richard S. Kleber, St. Olaf; Loren C. Larson, St. Olaf; John G. Lewis, St. Olaf; Kenneth O. May, University of Toronto; R.W. Nau, Carleton; William C. Ramaley, Carleton; J. Arthur Seebach, Jr., St. Olaf; Linda A. Seebach, St. Olaf; T.A. Vessey, St. Olaf; Frank L. Wolf, Carleton.

NEWS AND NOTICES

EDITED BY RAOUL HAILPERN, SUNY at Buffalo

Readers are invited to contribute to the general interest of this Department by sending news items to Mathematical Association of America, 1225 Connecticut Avenue, N.W., Washington, D.C. 20036. Items must be submitted at least two months before publication can take place.

PERSONAL ITEMS

Professor R. B. Deal, Jr., The University of Oklahoma, Medical Center, represented the Association at the inauguration of Dolphus Whitten, Jr., as President of Oklahoma City University on September 21, 1970.

University of Washington: Associate Professor E. B. Curtis, M.I.T., has been appointed Associate Professor; Professor M. E. Mahowald, Northwestern University, has been appointed Visiting Professor; Dr. A. W. Marshall, Boeing Scientific Research Laboratories, has been appointed Visiting Professor; Professor Paul Olum, Cornell University, has been appointed Visiting Professor; Associate Professor Jack Segal has been promoted to Professor.

University of Wisconsin, Milwaukee: Assistant Professors R. L. Gantos and D. W. Solomon have been promoted to Associate Professors.

Assistant Professor T. F. Banchoff, Brown University, has been promoted to Associate Professor.

Assistant Professor W. D. McIntosh, University of Missouri, Columbia, has been appointed Professor and Chairman of the Mathematics Department at Central Methodist College.

SECOND INTERNATIONAL CONGRESS ON MATHEMATICAL EDUCATION

The Second International Congress on Mathematical Education is scheduled to be held in Exeter, England, August 29–September 2, 1972. The Congress program will include major speakers in plenary sessions, a number of Working Groups covering a wide range of special projects and topics in mathematics education, and various other activities. Interested persons wishing to receive *Notices* from the Congress organizing committee may write to: D. G. Crawford, Honorary Secretary, I.C.M.I. Congress, Department of Education, University of Exeter, Thornlea, New North Road, Exeter EX4 4JZ, Devon, England.

MATHEMATICAL ASSOCIATION OF AMERICA

Official Reports and Communications

PRICES OF CARUS MONOGRAPHS AND MAA STUDIES

Effective June 1, 1971 the list prices of CARUS MONOGRAPHS and MAA STUDIES IN MATHEMATICS will be \$8.00. Members of the Association will continue to have the privilege of purchasing one copy of each of these books at half the list price, \$4.00. The officers of the Association regret that rising costs of printing and handling have made this price increase necessary in order to avoid significant subsidies from membership dues.

Members should continue to order single copies at the special price through the Washington Office. All sales at the list price are made directly by the distributors: The Open Court Publishing Company, La Salle, IL 61301 for CARUS MONOGRAPHS 1–4 and 6–8; John Wiley and Sons, 605 Third Avenue, New York, NY 10016 for CARUS MONOGRAPHS 9–15; Prentice-Hall, Inc., Englewood Cliffs, NJ 07631 for the MAA STUDIES.

CALENDAR OF FUTURE MEETINGS

Fifty-second Summer Meeting, Pennsylvania State University, University Park, August 30–September 1, 1971.

Fifty-fifth Annual Meeting, Las Vegas, Nevada, January 19–21, 1972.

The following is a list of the Sections of the Association with dates of future meetings so far as they have been reported to the Editorial Director.

ALLEGHENY MOUNTAIN
FLORIDA
ILLINOIS
INDIANA
IOWA
KANSAS
KENTUCKY
LOUISIANA-MISSISSIPPI
MARYLAND-DISTRICT OF COLUMBIA-VIRGINIA
METROPOLITAN NEW YORK
MICHIGAN
MISSOURI
NEBRASKA
NEW JERSEY
NORTH CENTRAL
NORTHEASTERN, Colby College, Waterville, Maine, June 19, 1971.
NORTHERN CALIFORNIA
OHIO
OKLAHOMA-ARKANSAS
PACIFIC NORTHWEST, Oregon State University, Corvallis, June 18–19, 1971.
PHILADELPHIA, Lafayette College, Easton, Pennsylvania, November 20, 1971.
ROCKY MOUNTAIN
SOUTHEASTERN
SOUTHERN CALIFORNIA
SOUTHWESTERN
TEXAS
UPPER NEW YORK STATE
WISCONSIN

FUTURE MEETINGS OF OTHER ORGANIZATIONS

AMERICAN ASSOCIATION FOR THE ADVANCEMENT OF SCIENCE, Philadelphia, December 26–31, 1971.

AMERICAN MATHEMATICAL SOCIETY, Pennsylvania State University, University Park, August 31–September 3, 1971.

AMERICAN SOCIETY FOR ENGINEERING EDUCATION, U. S. Naval Academy, Annapolis, June 21–24, 1971.

ASSOCIATION FOR COMPUTING MACHINERY, Chicago, August 3–5, 1971.

ASSOCIATION FOR SYMBOLIC LOGIC, Universidad Católica de Chile, Santiago, July 26–August 1, 1971.

CENTRAL ASSOCIATION OF SCIENCE AND MATHEMATICS TEACHERS, Detroit, Michigan, November 18–20, 1971.

FIBONACCI ASSOCIATION, College of the Holy Names, Oakland, California, November 13, 1971.

INSTITUTE OF MATHEMATICAL STATISTICS, Fort Collins, Colorado, August 23–26, 1971.

MU ALPHA THETA, Pennsylvania State University, University Park, September 1, 1971.

NATIONAL COUNCIL OF TEACHERS OF MATHEMATICS, Chicago, Illinois, April 16–20, 1972.

OPERATIONS RESEARCH SOCIETY OF AMERICA, Disneyland Hotel, Los Angeles, October 27–29, 1971.

PI MU EPSILON, Pennsylvania State University, University Park, August 31–September 1, 1971.

SOCIETY FOR INDUSTRIAL AND APPLIED MATHEMATICS, University of Washington, Seattle, Washington, June 28–30, 1971.

1971 GROUP FLIGHTS TO EUROPE

For members of the Mathematical Association of America and their families on regularly scheduled commercial jet flights. All fares are quoted round-trip, and are approximately 50% below normal fares.

June 7	PAN AM	New York-London	return July 5	$292.00
June 8	BOAC	New York-London	return Sept. 8	$292.00
June 23	TWA	New York-Paris	return Aug. 30	$307.00
June 28	TWA	New York-London	return Aug. 30	$292.00
July 6	TWA	New York-Paris	return Aug. 26	$307.00
July 13	TWA	New York-Paris	return Aug. 25	$307.00
Aug. 3	PAN AM	New York-London	return Sept. 2	$292.00
Aug. 10	BOAC	New York-London	return Sept. 9	$292.00

All fares, as quoted, include the $5.00 international transportation tax.

FLIGHT INFORMATION

Above fares are quoted for adult bookings. When computing for half-fare, first deduct the $5.00 tax, divide the fare in half and then add the $5.00 tax to this amount. For infant fares deduct the $5.00 tax, take 10% of the fare and then add the $5.00 tax to this amount.

Children under the age of two years, carried by the member, must pay 10% of the fare. Children from two years until their twelfth birthday must pay 50% of the fare. Children twelve years and older must pay full fare. The ages of the children and their names MUST be supplied with the application form and be computed as of the day of departure and not the day of reservation.

The flights listed are group flights on regularly scheduled commercial airlines at a discount of about 50% over the normal fares. These are NOT charters. The flights are open to those who are members at least six months prior to departure of the flight and to their spouse, dependent children and parents residing in the same household. No other persons are eligible for the group rate. Other relatives, friends, or students can travel on these flights with members, but only at the regular economy fare in effect for the flight. Airline regulations require that the member accompany the family members on the flight if they are to benefit from the group rate.

Reservations should be returned with a deposit of $100.00 for each full-fare reservation, and $50.00 for each half-fare. The balance of the payment is due by 90 days prior to departure. Members who book on flights within 90 days of departure should make full payment to reserve seats on the flight of their choice.

RESERVATIONS All bookings are accepted in the order they arrive at the office of the National Center for Educational Travel, Inc. Each booking is based on round-trips only. The airlines do not permit one-way or splitting of reservations. We encourage you to make early reservations in order to be certain of the availability of seats on the desired flight. It is necessary to return unreserved seats to the airlines thirty days before departure. If fewer than 40 seats are reserved on any flight, then thirty days before departure each person who has made a reservation on that particular flight will be notified of the cancellation of the group flight fare. During the past years, very few such cancellations have been necessary. However, if it is necessary to do so, it may be possible to travel on the same flight at the economy fare or to make reservations on a different flight under the Group Flights program.

Should you be interested in taking a trip to Europe this summer, but would prefer something of shorter duration and more flexibility as to dates, we would be more than happy to book you on a 29-45 day excursion fare trip. Round-trip fare to London (from New York and return) is $337.00. The round-trip fare to Paris is $362.00. A weekend surcharge of $15 each way will apply to excursion fares when transatlantic travel is on a Friday, Saturday or Sunday. Please contact NCET for further information.

CANCELLATIONS In case it is necessary for you to cancel your reservations, the following cancellation fee schedule has been established: (1) $10.00 each (maximum of $25.00 per family) for cancellations made from the time of reservation until 30 days prior to departure; (2) For the period of 30 to 21 days before departure, the airlines charge a 10% cancellation fee against each reservation; (3) For the period between 21 days and departure, the airlines charge a 25% cancellation fee against each reservation.

WAIVER The MAA, its officers, employees, staff, and agent assume no responsibility or liability in connection with these flights for any reason. A waiver is required of all participants. Please return the waiver with your reservation application.

Anyone interested in the above flights should write directly to: **National Center for Educational Travel, 1601 Connecticut Avenue, N.W., Washington, D.C. 20009, Telephone: (Area Code 202) 232-1456.**

GEORGE BANTA COMPANY, INC., MENASHA, WISCONSIN

THE AMERICAN MATHEMATICAL MONTHLY

(FOUNDED IN 1894 BY BENJAMIN F. FINKEL)

THE OFFICIAL JOURNAL OF

THE MATHEMATICAL ASSOCIATION OF AMERICA, INC.

VOLUME 78 NUMBER 6

CONTENTS

(*Continued on inside cover*)

JUNE-JULY 1971

NOTICE TO AUTHORS

Specialized research is usually unsuitable; see *Statement of Policy* (vol. 76, p. 2). Manuscript preparation: Please use the *Manual for Monthly Authors* (vol. 78, p. 1) and follow the format in current issues of the Monthly. Manuscripts should be typewritten, triple-spaced with wide margins; submit two copies and keep one for protection against loss. Each author receives 50 free reprints.

Backlog: Main Articles 9 months, Math. Notes 9 months, Research Problems 7 months, Classroom Notes 8 months, Math. Education 6 months.

EDITORIAL CORRESPONDENCE AND MAIN ARTICLES: to Harley Flanders, American Mathematical Monthly, Tel Aviv University, Ramat Aviv, Israel (see Notice, vol. 77, 1970, p. 555); NOTES, etc.: to the corresponding Associate Editor; ADVERTISING CORRESPONDENCE: to Raoul Hailpern, Mathematical Association of America, SUNY at Buffalo, Buffalo, N. Y. 14214; CHANGE OF ADDRESS and SUBSCRIPTIONS: to A. B. Willcox, Mathematical Association of America, 1225 Connecticut Ave., N.W., Washington, D. C. 20036.

Annual dues for members of the Association (including a subscription to the American Mathematical Monthly) are $12.50. For nonmembers the subscription price is $18.00.

Published by the Association at Washington, D. C., and Menasha, Wisconsin during the months of January, February, March, April, May, June–July, August–September, October, November, December.

Second-class postage paid at Washington, D. C., and additional mailing offices. Acceptance for mailing at special rate of postage provided for in the Act of February 28, 1925, embodied in Paragraph 4, Section 538, P. L. and R., authorized April 1, 1926.

PRINTED IN THE UNITED STATES OF AMERICA

FABER POLYNOMIALS AND THE FABER SERIES

J. H. CURTISS, University of Miami

1. Introduction. A problem in complex analysis which attracted the attention of a number of distinguished mathematicians around the turn of the century was that of finding a set of polynomials, $p_1(z)$, $p_2(z)$, $\cdots$, which "belong" to a given region, in the sense that any function f analytic in the region can be expanded in a convergent series $a_0+\sum_1^\infty a_j p_j(z)$, in which the coefficients a_j, but not the polynomials p_j, depend on f. In 1903, Georg Faber [9] published a solution to the problem which was notable both for the basic simplicity of the convergence proof and also for the rich and interesting structure of the polynomials.

Over the years a voluminous literature of research papers concerned with the Faber polynomials has appeared. Successful applications of the polynomials have been made in the following problem areas, among others: best polynomial approximation in the Chebychev sense (Faber [10], Sewell [21]); degree of polynomial approximation (Sewell [21], Al'per [2]); asymptotic properties of polynomials orthogonal on a simple closed curve (Szegö [24]); solutions of the Dirichlet problem by harmonic polynomial interpolation (Curtiss [7], [8]); necessary and sufficient conditions for an analytic function to be univalent (Grunsky [11]); extremum problems in conformal mapping (Schiffer [19]); the coefficient problem in conformal mapping (Jenkins [13]), (Charzynski and Schiffer [5]); estimates of the discriminant of a plane continuum (Pommerenke [15], [16]); and the distribution of certain extremal point systems (the Fekete points) on a continuum (Pommerenke [17]). Many of these papers, particularly those of Pommerenke, contain substantial contributions to the available information about the structure of the Faber polynomials. We do not pretend that this is a complete guide to the theory and applications of the Faber polynomials. A further compilation of references, many in the Russian language, will be found in the expository paper of Suetin [23], itself in the Russian language.

The purposes of this paper are (a) to give an exposition of the structure of the Faber polynomials from what might be called a modern-classical point of

J. H. Curtiss received his Ph.D. degree in mathematics at Harvard University in 1935. His thesis was on complex approximation theory and was written under the direction of J. L. Walsh. After a year as instructor at Johns Hopkins University, he joined the faculty of Cornell University where he was first an instructor and then an assistant professor. He served as an officer in the Naval Reserve in 1943–46, with terminal rank of Lieutenant Commander. From 1946–1953 he was Assistant to the Director and then Chief of the Applied Mathematics Laboratories at the National Bureau of Standards. After a year as a senior scientist at the Courant Institute of New York in 1953–54, he was appointed Executive Director of the American Mathematical Society, where he served until 1959. Since September 1959 he has held a professorship in mathematics at the University of Miami, Coral Gables, Florida. He is a fellow of the American Association for the Advancement of Science, of the Institute of Mathematical Statistics, and of the American Statistical Association. He is a past president of the Association for Computing Machinery and a past vice-president of the Institute of Mathematical Statistics. His principal research interests are approximation by complex and harmonic polynomials and applied probability theory. *Editor.*

view, using Lebesgue integration theory where warranted, and (b) to present, as a by-product, a modest contribution to the convergence of the Faber series which may suggest new questions to explore. The convergence study, which is all in the last section, involves a connection between the Faber series and a related Fourier series (the lemma in Section 5) which may have been hitherto overlooked. The background expected of the reader can be roughly described as that provided by a two-semester course in real analysis and a two-semester course in complex analysis, but certain standard results in modern Fourier series theory not usually included in such courses will be needed. Specific references for these results will be cited, all in the wonderfully comprehensive monograph by Zygmund [25].

2. The Faber polynomials; formal relations. Let E be a compact set of the complex plane with a complement E^C which is simply connected in the extended complex plane. According to the Riemann mapping theorem, there exists a function

$$z = \phi(w) = d\left[w + d_0 + \frac{d_1}{w} + \frac{d_2}{w^2} + \cdots\right], \tag{2.1}$$

(where $d>0$ is the transfinite diameter or capacity of E), which is univalent and analytic for $|w|>1$, and which maps E^C conformally onto $\{w: |w|>1\}$. For z exterior to a sufficiently large circle, the inverse function $\phi^{-1}(z)$ exists and has a Laurent expansion of the following form:

$$\phi^{-1}(z) = \frac{z}{d} + g_0 + \frac{g_1}{z} + \frac{g_2}{z^2} + \cdots .$$

The nth Faber polynomial $p_n(z)$, $n=1, 2, \cdots$, belonging to E (or to ϕ) is the part of the Laurent series for $[\phi^{-1}(z)]^n$ which contains the nonnegative powers of z—the "principal part at infinity" of this Laurent series. Clearly this is a polynomial of exact degree n with leading term $(z/d)^n$.

Let $C_R = \{z: z=\phi(s),\ |s| = R>1\}$. Consider the integrals, with $t=\phi(s)$,

$$\frac{1}{2\pi i}\int_{C_R} \frac{[\phi^{-1}(t)]^n}{t-z}\,dt = \frac{1}{2\pi i}\int_{|w|=R} \frac{s^n\phi'(s)ds}{\phi(s)-z}$$

for $z\in \operatorname{Int} C_R$. The path of the integral on the left can be replaced by a circle of radius large enough so that $\phi^{-1}(z)$, and therefore $[\phi^{-1}(z)]^n$, has a uniformly convergent Laurent series on this circle. The Laurent series can be integrated term by term, and when this is done the integral reproduces the principal part of the series at infinity and kills off the terms with negative exponents. Thus

$$p_n(z) = \frac{1}{2\pi i}\int_{|s|=R} \frac{s^n\phi'(s)ds}{\phi(s)-z}, \qquad n = 1, 2, 3, \cdots, z\in \operatorname{Int} C_R. \tag{2.2}$$

Now with z fixed on Int C_R, the function $s\phi'(s)/[\phi(s)-z]$ is analytic for

$|s| \geqq R$ and has the value 1 at $s=\infty$, so it has a Laurent series

$$\frac{s\phi'(s)}{\phi(s)-z} = 1 + p_1\frac{1}{s} + p_2\frac{1}{s^2} + \cdots, \qquad |s| \geqq R,\ z \in \operatorname{Int} C_R. \tag{2.3}$$

But the Cauchy formulas for the coefficients in (2.3) are precisely the integrals appearing in (2.2), so these coefficients p_n are indeed $p_n(z)$ for each n. We thus have a generating function for the Faber polynomials. Another generating function which appears in the literature (e.g., [19]) is obtained by dividing both sides of (2.3) by s and taking indefinite integrals term by term. The result after an appropriate choice of the constant of integration is

$$\ln\left[\frac{\phi(s)-z}{sd}\right] = -\sum_{n=1}^{\infty}\frac{1}{n}\,p_n(z)\,\frac{1}{s^n}, \qquad z \in \operatorname{Int} C_R,$$

where here, as in the sequel, a branch of logarithm is used for which $\ln 1 = 0$.

A recursion formula for the Faber polynomials is easily derived. We multiply both sides of (2.3) by $\phi(s)-z$ and expand $\phi(s)$ and $\phi'(s)$ in their Laurent series. By comparing coefficients of like powers of s, we obtain

$$p_1(z) = \frac{z}{d} - d_0, \tag{2.4}$$

$$\begin{aligned} p_{n+1}(z) &= p_1(z)p_n(z) - d_1p_{n-1}(z) - d_2p_{n-2}(z) - \cdots \\ &\quad - d_{n-1}p_1(z) - (n+1)d_n, \qquad n = 1, 2, \cdots. \end{aligned} \tag{2.5}$$

(In using this, we let $p_0 = p_{-1} = p_{-2} = \cdots = 0$.)

We shall now examine $p_n(\phi(w)) = F_n(w)$. Choose R_1 so that $1 < R_1 < R$ (where R appears in (2.2)), and let $z = \phi(w)$ lie in the region bounded by C_{R_1} and C_R. The function $s^n\phi'(s)/[\phi(s)-\phi(w)]$ as a function of s is analytic in the closed region except for a simple pole at $s=w$. The residue at the pole is

$$\lim_{s\to w}(s-w)\frac{s^n\phi'(s)}{\phi(s)-\phi(w)} = w^n.$$

Thus by the residue theorem,

$$p_n(z) = F_n(w) = w^n + \frac{1}{2\pi i}\int_{|s|=R_1}\frac{s^n\phi'(s)ds}{\phi(s)-\phi(w)}. \tag{2.6}$$

But the integral on the right is an analytic function of w for $|w| > R_1$ and has the value 0 at $w=\infty$. Thus it has a Laurent series in w convergent at least for $|w| > R$. We write this series in the form

$$F_n(w) = w^n + \sum_{k=1}^{\infty}\alpha_{nk}w^{-k}. \tag{2.7}$$

The series converges for all w, $|w| > 1$, uniformly for $|w| > R_1 > 1$, where R_1 is otherwise arbitrary, and

$$(2.8)\qquad \alpha_{nk} = -\frac{1}{4\pi^2}\int_{|w|=R}\int_{|s|=R_1}\frac{w^{k-1}s^n\phi'(s)dsdw}{\phi(s)-\phi(w)}.$$

We shall call the numbers α_{nk} the Faber coefficients of E. It should be noted that d has disappeared from the scene; the α_{nk}'s are independent of the value of d in the sense that if d is varied in (2.1) and the coefficients $d_1, d_2, \cdots$ are held constant, the α_{nk}'s remain constant.

Given any sequence of polynomials $P_1(z), P_2(z), \cdots$ in which P_n is of exact degree n for $n=1, 2, \cdots$, it is easy to show by induction that the additional condition that $P_n(\phi(w))$ shall have a Laurent series of the form $w^n+\sum_{k=1}^{\infty}\beta_{nk}w^{-k}$, $|w|>1$, $n=1, 2, \cdots$, uniquely determines the coefficients of each P_n. Therefore with this condition it must be true that $P_n(z)$ is identically equal to $p_n(z)$, the nth Faber polynomial belonging to ϕ. This remark provides an alternative definition of the Faber polynomials which often appears in the literature.

The Faber coefficients have a generating function of their own, from which an interesting law of symmetry emerges. We take s and w with $|s|>|w|>1$. Now from (2.3) and (2.7)

$$\begin{aligned}\frac{\partial}{\partial s}\ln\frac{\phi(s)-\phi(w)}{d(s-w)} &= \frac{\phi'(s)}{\phi(s)-\phi(w)}-\frac{1}{s-w}\\ (2.9)\qquad &= \frac{1}{s}+F_1(w)\frac{1}{s^2}+F_2(w)\frac{1}{s^3}+\cdots-\frac{1}{s}-\frac{w}{s^2}-\frac{w^2}{s^3}-\cdots\\ &= \sum_{n=1}^{\infty}(F_n(w)-w^n)s^{-n-1} = \sum_{n=1}^{\infty}\left(\sum_{k=1}^{\infty}\alpha_{nk}w^{-k}\right)s^{-(n+1)}.\end{aligned}$$

For any fixed w, $|w|>1$, this power series in s converges uniformly and absolutely for $|s|\geqq R_1>|w|$. The analytic function of s defined by

$$-\sum_{n=1}^{\infty}n^{-1}\left(\sum_{k=1}^{\infty}\alpha_{nk}w^{-k}\right)s^{-n},\qquad |s|\geqq R_1,$$

has a derivative with respect to s which is identically equal to the last member of (2.9), and so this function must differ from $\ln[\phi(s)-\phi(w)]/(w-s)d$ by only a constant depending on w. As usual choose a branch of the logarithmic function for which $\ln 1=0$. By letting $s\to\infty$ in (2.9) we see that the constant is zero. Thus

$$(2.10)\qquad \ln\frac{\phi(s)-\phi(w)}{(s-w)d} = -\sum_{n=1}^{\infty}\sum_{k=1}^{\infty}\frac{\alpha_{nk}}{n}w^{-k}s^{-n}.$$

The function $[\phi(s)-\phi(w)]/(s-w)$ can never vanish in the Cartesian product domain $\{s:|s|>1\}\times\{w:|w|>1\}$, because ϕ is univalent and $\phi'(w)\neq 0$ for all w, $|w|>1$. It is an analytic function of two complex variables in this domain. Thus the left member of (2.10) is also an analytic function of s and w in this domain, and the right member of (2.10) must be its power series expansion

about (∞, ∞). It follows that the equation (2.10) is valid for all (s, w), $|s|>1$, $|w|>1$. In particular, the restriction $|s|>|w|$ used to derive (2.9) and (2.10) can be ignored in using these formulas.

The left side of (2.10) is symmetric in s and w, and so these variables can be interchanged without changing the value of the function. It follows that $\alpha_{nk}/n = \alpha_{kn}/k$, or $k\alpha_{nk} = n\alpha_{kn}$, $n, k = 1, 2, \cdots$. This is the Grunsky Law of Symmetry [11].

We now derive a formula involving a summation over the squared absolute values of all the Faber coefficients. This will be used later in Section 5 in a discussion of the Faber series.

Given any function $g(s)$, meromorphic for $|s|>1$ with the only pole at $s=\infty$, the Laurent series for $g(s)$ has the appearance $\sum_{k=-m}^{\infty} a_k s^{-k}$. This converges absolutely on any circle $\{s: |s| = R > 1\}$. Therefore the Cauchy product series

$$\sum_{n=-m}^{\infty} \left(\sum_{k=-m}^{n} a_k \bar{a}_{n-k} s^{-k} \bar{s}^{-(n-k)} \right) = \sum_{n=-m}^{\infty} a_n s^{-n} \sum_{n=-m}^{\infty} \bar{a}_n \bar{s}^{-n}$$

converges absolutely and uniformly in s to the function $g(s)\overline{g(s)} = |g(s)|^2$, where the bar denotes complex conjugate. (See [12], vol. I, p. 113.) Integration of the product series term by term is thereby justified, and after a simple computation, we obtain, with $s = R\exp(i\sigma)$:

$$\text{(2.11)} \qquad \frac{1}{2\pi}\int_0^{2\pi} |g(s)|^2 d\sigma = \sum_{n=-m}^{\infty} |a_n|^2 R^{-2n}.$$

Applying the formula to (2.9), again with $s = R\exp(i\sigma)$ and with w fixed on $\{w: |w|>1\}$, we obtain

$$\frac{1}{2\pi}\int_0^{2\pi} \left| \frac{\partial}{\partial s} \ln \frac{\phi(s) - \phi(w)}{s - w} \right|^2 d\sigma = \sum_{n=1}^{\infty} \left[R^{-2n-2} \left| \sum_{k=1}^{\infty} \alpha_{nk} w^{-k} \right|^2 \right].$$

The integral on the left side represents a continuous function of w for $|w|>1$, and so is integrable over $\{w: |w| = R\}$. The inner summation on the right side represents the analytic function $F_n(w) - w^n$, so for each n it is integrable on $|w| = R$. The overall series indexed by n on the right side is a series of positive integrable terms, and so by the Lebesgue monotone convergence theorem the series can be integrated term by term with respect to w, $|w| = R$, and the resulting series will represent the integral over $|w| = R$ of the left member. We proceed to carry out this integration and use (2.11) again to evaluate the integrals on the right side. The result is

$$\text{(2.12)} \qquad \frac{1}{4\pi^2}\int_0^{2\pi}\int_0^{2\pi} \left| \frac{\partial}{\partial s} \ln \frac{\phi(s) - \phi(w)}{s - w} \right|^2 d\sigma d\theta = \sum_{n=1}^{\infty}\sum_{k=1}^{\infty} R^{-2n-2k-2} |\alpha_{nk}|^2,$$
$$s = Re^{i\sigma}, \quad w = Re^{i\theta}.$$

This is the desired formula.

There is a recursion formula for the Faber coefficients [7]. In the first place, from (2.4) with $z=\phi(w)$, we have $\alpha_{1k}=d_k$, $k=1, 2, \cdots$ (see (2.1) for d_k). Multiply both sides of (2.7) by $1/w$ and integrate over $|w|=R$. The result is zero. Multiply both sides of (2.5) by $1/2\pi iw$ with $z=\phi(w)$ and integrate over $|w|=R$. The left side vanishes, and on the right side there will be a residue at $w=0$ which turns out to be $\alpha_{n1}+d_n-(n+1)d_n$, so $\alpha_{n1}=nd_n$, $n=1, 2, \cdots$. Then substitute $z=\phi(w)$ into (2.5) and equate coefficients of w^{-k} on the left and right sides of (2.5). We obtain the recursion formula:

$$
\begin{aligned}
\alpha_{n,k+1} &= \alpha_{n+1,k} - d_1\alpha_{n,k-1} - d_2\alpha_{n,k-2} - \cdots - d_{k-1}\alpha_{n1} \\
&\quad + d_1\alpha_{n-1,k} + d_2\alpha_{n-2,k} + \cdots + d_{n-1}\alpha_{1,k} - d_{k+n}
\end{aligned}
\tag{2.13}
$$

with initial values $\alpha_{1k}=d_k$, $k=1, 2, \cdots$, $\alpha_{n1}=nd_n$, $n=1, 2, \cdots$. The formula is valid for all $n \geqq 1$, $k \geqq 1$ when each letter with zero or negative subscript is given the value zero.

Without referring specifically to this recursion relation, Schur [20] in effect solved it by deriving an explicit formula for the coefficient α_{nk} in term of the coefficients $d_1, d_2, \cdots$. It is fairly apparent from (2.13) that α_{nk} is a polynomial in the d_k's, but Schur discovered the interesting fact that the polynomial had *nonnegative integer* coefficients. A shorter proof was given by the author in [7] without, however, exhibiting the explicit formula for α_{nk}.

The special case in which the boundary ∂E of E is a simple closed *analytic* curve (see [21], pp. 226–227) deserves attention, if only for historical reasons. (Faber's construction [9] and many later results concerning his polynomials have been restricted to this situation.) Here the exterior mapping function ϕ can be extended into the disk $\{w: |w|<1\}$ so as to be analytic and univalent for $|w|>r_0$, with $0 \leqq r_0 < 1$. We henceforth take r_0 to be the least number with the indicated property and with Pommerenke [15] call ∂E an r_0-analytic curve. It is now possible to replace the condition $R>1$, $R_1>1$ on the level curves C_R, C_{R_1} in (2.2), (2.3), (2.6), (2.7), and (2.8), by $R>r_0$, $r_0<R_1<R$. The difference quotient $[\phi(s)-\phi(w)]/(s-w)$ is analytic and nonvanishing (because of the univalence of ϕ) for $(s, w) \in \{|s|>r_0\} \times \{|w|>r_0\}$, so (2.9) and (2.10) are now valid for all such (s, w). The R in (2.12) can be taken to be merely greater than r_0.

3. Examples. It would appear that the only specific geometric types of set E for which the Faber polynomials can be represented by reasonably simple explicit formulas are line segments, closed circular and elliptical disks, and lemniscates. We give three examples:

(A) *E is the closed disk $|z-a| \leqq d$.* Then

$$
\phi(w) = dw + a, \qquad \phi^{-1}(z) = (z-a)/d,
$$
$$
p_n(z) = [(z-a)/d]^n, \qquad p_n[\phi(w)] = F_n(w) = w^n.
$$

(B) *E is the closed region bounded by the ellipse* $(\operatorname{Re} z/a)^2+(\operatorname{Im} z/b)^2=1$, $a>b>0$. Let $c=(a^2-b^2)^{1/2}/2$, $P=[(a+b)/(a-b)]^{1/2}$, so $P>1$. Then $\phi(w)=cPw+c/Pw$, which is analytic and univalent for $|w|>1/P$. The first member of

(2.9) becomes $P^{-2}s^{-2}w^{-1}[1-P^{-2}s^{-1}w^{-1}]^{-1}$, so the coefficient of s^{-n-1} is $P^{-2n}w^{-n}$, and $F_n(w)=w^n+P^{-2n}w^{-n}$. Solving the mapping function $z=\phi(w)$ for w in terms of z and substituting, we obtain

$$p_n(z) = 2\sum_{k=0}^{[n/2]}\binom{n}{2k}z^{n-2k}(z^2-4c^2)^k, \qquad n=1,2,\cdots.$$

(C) *E is the closed region bounded by the three-cusped hypocycloid with parametric equation* $z=2\exp i\theta+\exp(-2i\theta)$, $0\leqq\theta\leqq 2\pi$. Then $\phi(w)=2w+w^{-2}$. Even in this relatively simple case it is not easy to sort out the coefficients in the generating functions (2.3) and (2.9) or (2.10). From the recurrence (2.4) and (2.5), with $t=z/d=z/2$, $z=2t$, we obtain

$$p_1(2t)=t,\quad p_2(2t)=t^2,\quad p_3(2t)=t^3-3,\quad p_4(2t)=t^4-t,$$
$$p_5(2t)=t^5-5,\quad p_6(2t)=t^6-6t^3+3,\quad p_7(2t)=t^7-5t^4+7.$$

The general case is obtainable by solving the third-order difference equation $p_{n+3}-tp_{n+2}+p_n=0$.

Faber in [9], using the generating function (2.3), derived the formula for $p_n(z)$, in the case in which E is bounded by the lemniscate $\{z:|z^2-1|=1\}$.

4. Estimates and convergence theorems for the Faber polynomials. The results to be considered here can be roughly placed in three classes: (a) estimates related to the analyticity of the exterior mapping function ϕ; (b) facts arising from the Fourier character of the Faber coefficients; (c) inequalities which stem from the Grunsky inequalities, to be described below, and which are related to the Gronwall Area Theorem.

Class (*a*): Two obvious inequalities can be obtained by applying the Cauchy coefficient estimates to (2.6) and (2.8). In (2.6) let $|w|=\rho$, $R_1<\rho<R$. We obtain

$$|p_n(z)| = |F_n(w)| \leqq \rho^n + R_1^{n+1}\max\left[\left|\frac{\phi'(s)}{\phi(s)-\phi(w)}\right|, \ |s|=R_1, \ |w|=\rho\right],$$

which can be abbreviated to

$$|p_n(z)| = |F_n(w)| \leqq M\rho^n, \qquad n=1,2,\cdots, \quad z=\phi(w), \quad |w|=\rho, \tag{4.1}$$

where M depends on ρ and R_1. By the Maximum Modulus Principle the inequality with deletion of the middle member remains valid for all $z\in\operatorname{Int} C_\rho$. The same technique applied to (2.8) with R replaced there by ρ yields

$$|\alpha_{nk}| \leqq M\rho^{n+k+1}, \qquad k, n = 1,2,\cdots. \tag{4.2}$$

These inequalities are chiefly of interest in the analytic-boundary case described at the end of Section 4.2, because then ρ can be taken to be less than one. For the general case Smirnov and Lebedev in [22], Chapter 2, give more elaborate versions in which the M is replaced by explicit formulas involving, in our notation, ρ and R_1.

It will be shown below that when the boundary ∂E is an r_0-analytic curve, the following improved version of (4.2) can be derived by considerations to be taken up under Class (c):

$$|\alpha_{nk}| \leqq \left(\frac{n}{k}\right)^{1/2} r_0^{n+k}. \tag{4.3}$$

Sharp estimates for the modulus $|p_n(z)|$, $z \in E$, based on the analyticity of $\phi(w)$ when E is any compact set with simply connected E^C with capacity $d=1$ (see (2.1)) have recently been given by Kövari and Pommerenke [14]. A typical result is this:

THEOREM 4.1. (i) *There exist absolute constants A and $\alpha<1/2$ such that* $\max_{z\in E}|p_n(z)| \leqq An^\alpha$. (ii) *There exists an exterior mapping function ϕ such that for each fixed z, the associated $p_n(z)$ satisfies $|p_n(z)|>n^{0.138}$ for an infinite sequence of n's.*

Only the crude estimates (4.2) are needed to develop from (2.7) a pair of classical asymptotic formulas. Take any w with $|w|>1$, or with $|w|>r_0$ in the analytic-boundary case. It is easy to show by using (4.2) that

$$\lim_{n\to\infty} w^{-n}\sum_{k=1}^{\infty}\alpha_{nk}w^{-k} = 0.$$

Thus with $z=\phi(w)$ we have

$$\lim_{n\to\infty}\frac{p_{n+1}(z)}{p_n(z)} = \lim_{n\to\infty}\frac{F_{n+1}(w)}{F_n(w)} = \lim_{n\to\infty}\frac{w+w^{-n}\sum_{k=1}^{\infty}\alpha_{n+1,k}w^{-k}}{1+w^{-n}\sum_{k=1}^{\infty}\alpha_{n,k}w^{-k}} = w = \phi^{-1}(z),$$

and

$$\lim_{n\to\infty}|p_n(z)|^{1/k} = |w| = |\phi^{-1}(z)|,$$

where, as in section 2, ϕ^{-1} is the inverse of ϕ.

Class (*b*): A basic proposition concerning the Fourier character of the Laurent series, (2.7) for $F_n(w)$ is the following:

THEOREM 4.2. *Let E be a compact set with E^C simply connected in the extended plane. Then for each n, $n=1, 2, \cdots$, $\lim_{R\downarrow 1} F_n(R\exp(i\theta)) = F_n(\exp i\theta)$ exists almost everywhere on $\{\theta: 0\leqq\theta\leqq 2\pi\}$; $F_n(\exp(i\theta))$ is integrable in the sense of Lebesgue and uniformly bounded in θ (but not necessarily in n); and the Laurent series (2.7) for $F_n(w)$ with $w=\exp(i\theta)$ is the formal trigonometric Fourier series for $F_n(\exp(i\theta))$.*

("Almost everywhere" means at all points except possibly on a set of Lebesgue measure zero—abbreviated to a.e. The formal trigonometric Fourier series for

an integrable function $g(\theta)$ is the series $\sum_{k=-\infty}^{\infty} c_k \exp(ik\theta)$, where $c_k = (2\pi)^{-1}\int_0^{2\pi} g(\tau)$ $\cdot\exp(-ik\tau)d\tau$. Henceforth "integrable" will always mean absolutely integrable with respect to linear Lebesgue measure.)

For the proof, we first show that the theorem is true for $\phi(w)$, which, after all, is only a slightly modified version of $F_1(w)$. A chain of reasoning leading to this result runs as follows: Take any $R>1$; ϕ maps $\{w:1<|w|<R\}$ onto $\operatorname{Int} C_R \cap E^C$. Let $\Delta = \max|z|$, $z \in C_R$. The function with values $\phi(w)-wd$ is analytic for $|w| \geqq R$ including $w=\infty$, and so its absolute value takes on a maximum for all w, $|w| \geqq R$, on the circle $\{w:|w|=R\}$ (Maximum Modulus Principle). This maximum can be no greater than $\Delta + Rd$. On the other hand, for w in the annulus $\{w:1<|w|<R\}$, we have $|z| = |\phi(w)| < \Delta$. Thus $\phi(w)-wd$ is analytic and uniformly bounded for $|w|>1$, and $\phi(1/\zeta)-d/\zeta$ is analytic and uniformly bounded for $|\zeta|<1$. By Fatou's theorem ([12], vol. II, p. 364), there exist radially approached boundary values for $\phi(1/\zeta)-d/\zeta$ a.e. on $|\zeta|=1$, and therefore $\lim_{R\downarrow 1}\phi(R\exp i\theta) = \phi(\exp i\theta)$ exists a.e. The limit values are clearly bounded; they are measurable and therefore integrable. By the Lebesgue Bounded Convergence Theorem, the Laurent coefficients dd_j can be represented as follows:

$$
\begin{aligned}
dd_j &= \frac{1}{2\pi}\int_0^{2\pi} R^j e^{ij\theta}\phi(Re^{i\theta})d\theta = \lim_{R\downarrow 1}\frac{1}{2\pi}\int_0^{2\pi} R^j e^{ij\theta}\phi(Re^{i\theta})d\theta \\
&= \frac{1}{2\pi}\int_0^{2\pi} \lim_{R\downarrow 1}\left[R^j e^{ij\theta}\phi(Re^{i\theta})\right]d\theta = \frac{1}{2\pi}\int_0^{2\pi}\phi(e^{i\theta})e^{ij\theta}d\theta.
\end{aligned} \tag{4.4}
$$

The coefficients dd_j are therefore the Fourier coefficients of ϕ (exp $i\theta$). Thus the theorem is true for ϕ.

But $F_n(w) = p_n(z)$, with $z=\phi(w)$, is merely a polynomial of degree n in $\phi(w)$ with constant coefficients. Therefore it too must be bounded in any annulus $\{w:1<|w|<R\}$; and reflecting the behavior of ϕ, it must have bounded limit values for radial approach to $|w|=1$ from outside. The computation (4.4) is therefore valid with ϕ replaced by F_n, and dd_j by α_{nj}. This proves the theorem.

An endless vista of results concerning the behavior of the Laurent series (2.7) for F_n on $|w|=1$ is thereby opened up, corresponding to the many known results on Fourier series of bounded functions. Whenever a hypothesis on $\phi(w)$ for $|w| \geqq 1$ or $|w|=1$ can be carried over to a polynomial in $\phi(w)$, the corresponding conclusion for the Laurent-Fourier series (2.1) is valid for (2.7). Most of the classical convergence and summability criteria for the Fourier series of $\phi(\exp i\theta)$ are of this character; see [25], vol. I, in particular Chapters II and III. We shall take space here to present only two specific results of this category. The first is of interest because of the simplicity of its statement (but the known proof is very difficult) and the recency of its discovery. The second will be applied in the convergence theory of the Faber series which will be taken up later on in this paper.

In 1966, Lennart Carleson [4] published solutions of various hitherto unre-

solved problems in trigonometric Fourier series theory. In particular, he showed that the Fourier series of any function f with f^2 integrable converges almost everywhere to the value of the function. Our $F_n(\exp(i\theta))$ is bounded and measurable and so satisfies Carleson's hypothesis. We obtain

THEOREM 4.3. *For each n, $n=1, 2, \cdots$, the series* (2.7) *with $w=\exp(i\theta)$ converges almost everywhere on $[0, 2\pi]$ to $F_n(\exp(i\theta))$.*

For the second result we assume that the boundary ∂E of E is a simple closed curve. It is known that in this case ϕ can be extended so as to be a continuous function for $1 \leqq |w| < \infty$, and the correspondence between ∂E and $\{w: |w| = 1\}$ is one-to-one and bicontinuous (see [12] vol. II, p. 367, [25] vol. I, pp. 290 ff.). Further suppose that ∂E is rectifiable. Then $\phi(\exp(i\theta))$ is a continuous periodic function of bounded variation on $[0, 2\pi]$. (See [12], vol. I, p. 36.) It follows from the polynomial nature of $F_n(w)$ that $F_n(w)$ can also be extended so as to be continuous for $1 \leqq |w| < \infty$, and $F_n(\exp(i\theta))$ is of bounded variation. Then $F_n(w) - w^n$ is continuous for $1 \leqq |w| \leqq \infty$ and of bounded variation in θ with $w = \exp(i\theta)$. The series $\sum_{k=1}^{\infty} \alpha_{nk} \exp(-ki\theta)$ is the Fourier series for $F_n(\exp(i\theta)) - \exp in\theta$ (Theorem 4.2). Under these circumstances, the power series $\sum_{k=1}^{\infty} \alpha_{nk} w^{-k}$ becomes what is known as a power series of bounded variation, and a number of interesting facts are available. (See [25], vol. I, Chapter 7, Section 8.) The following theorem expresses two of the salient ones in the present context:

THEOREM 4.4. *If ∂E is a rectifiable simple closed curve, then for each n* (a) *$F_n(\exp(i\theta))$ is absolutely continuous and* (b) *the series* (2.7) *converges absolutely for $|w| = 1$.*

Then of course it converges absolutely for $|w| \geqq 1$.

Various theorems concerning the Faber coefficients α_{nk} can be derived from (2.9), (2.10), and (2.12) by imposing conditions on the mapping function ϕ which will insure more or less smooth boundary values for $[\phi(s) - \phi(w)]/(s-w)$ as $|s| \downarrow 1$, $|w| \downarrow 1$. A simple result of this type is obtainable immediately from (2.12):

THEOREM 4.5. *A necessary and sufficient condition for $\sum_{n=1}^{\infty} \sum_{k=1}^{\infty} |\alpha_{nk}|^2 < \infty$ is that there exists a constant $M > 0$ such that with $s = R\exp(i\sigma)$, $w = R\exp(i\theta)$,*

$$\int_0^{2\pi} \int_0^{2\pi} \left| \frac{\partial}{\partial s} \ln \frac{\phi(s) - \phi(w)}{s - w} \right|^2 d\sigma d\theta \leqq M \tag{4.5}$$

for all $R > 1$.

If (4.5) is valid, then the value of the double series is no greater than $(4\pi^2)^{-1}M$.

The functions $\phi'(w)$ and $\phi''(w)$ are analytic for $|w| > 1$, and if further they are bounded for $|w| > 1$, they have bounded limit functions $\phi'(\exp(i\theta))$, $\phi''(\exp i\theta)$ (Fatou). By the methods used in [6], Section 3, the following result can be proved:

THEOREM 4.6. *If ϕ' and ϕ'' are each bounded functions for $|w|>1$, and $\phi'(\exp i\theta)\neq 0$, $0\leqq\theta\leqq 2\pi$, then* (4.5) *is valid and so* $\sum\sum|\alpha_{nk}|^2<\infty$.

It is clear from (4.2) that if ∂E is an analytic simple closed curve, then $\sum\sum|\alpha_{nk}|^2<\infty$.

Class (*c*): The classical Area Theorem of Gronwall ([12], vol. II, p. 347) in the present notation takes the form $\sum_{n=1}^{\infty} n|d_n|^2\leqq 1$, where the d_n's are those which appear in (2.1). Pommerenke [15], by using a formula of Grunsky [11] for the area of a map given by a multivalent function, found the following simple generalization:

THEOREM 4.7. *Let $g(w)=b_{-N}w^N+\cdots+b_{-1}w^1+b_0+b_1w^{-1}+b_2w^{-2}+\cdots$ be analytic in $A=\{w:1<|w|<\infty\}$ and let it assume any of its values on at most N distinct points in A. Then*

$$\sum_{k=1}^{\infty} k|b_k|^2 \leqq \sum_{k=1}^{N} k|b_{-k}|^2.$$

Equality obtains if and only if $g(w)$ assumes each value exactly N times, except possibly for a set of values of Lebesgue planar measure zero.

The Area Theorem is the case in which $g(w)=\phi(w)$ and $N=1$.

To apply this to the Faber polynomials, first note that by the Fundamental Theorem of Algebra, $p_n(z)$ assumes any given value on at most n distinct points z. Since $\phi(w)$ is univalent, the function $p_n(\phi(w))=F_n(w)$, which is of the same type as the g in the theorem, assumes any one of its values on at most n points $w\in A$. From (2.7) and the theorem we obtain:

$$\sum_{k=1}^{\infty} k|\alpha_{nk}|^2 \leqq n, \qquad n=1,2,\cdots. \tag{4.6}$$

But Pommerenke [12] showed that we can do better than this, and at the same time get a sharper version of an inequality of Grunsky [11] derived by the latter, and by others later on, in more cumbersome ways. Choose any positive integer N and consider the polynomial $P(z)=\sum_{n=1}^{N} u_np_n(z)$, where the u_n's are arbitrary complex numbers, but not all zero. This polynomial is not a constant and is of degree at most N, and it therefore takes on any of its values on at most N points z. Since $\phi(w)$ is univalent, $P(\phi(w))$ takes on any value in at most N distinct points of the annulus A above. Now

$$P(\phi(w)) = \sum_{n=1}^{N} u_nF_n(w) = \sum_{n=1}^{N} u_nw^n + \sum_{n=1}^{N}\left[u_n\sum_{k=1}^{\infty}\alpha_{nk}w^{-k}\right], \qquad w\in A.$$

Hence, by identifying coefficients of powers of w and using Theorem 4.7, we obtain

$$\sum_{k=1}^{\infty} k\left|\sum_{n=1}^{N} u_n\alpha_{nk}\right|^2 \leqq \sum_{k=1}^{N} k|u_k|^2. \tag{4.7}$$

The inequality is trivially valid when all the u_n's are zero. We summarize:

THEOREM 4.8. (Pommerenke). *Let* $\langle u_n \rangle$ *be an arbitrary sequence of complex numbers. Then for each* N, $N=1, 2, \cdots$, (4.7) *and*

$$\sum_{n=1}^{\infty} \frac{1}{n} \left| \sum_{k=1}^{N} \alpha_{nk} u_k \right|^2 \leqq \sum_{k=1}^{N} \frac{1}{k} \, | u_k |^2 \tag{4.8}$$

are valid.

The inequality (4.8) is the version of (4.7) obtained by applying to (4.7) the Grunsky Law of Symmetry. From the last sentence of Theorem 4.7 it can be shown [15] that equality obtains here for all non-null sequences $\langle u_n \rangle$ if and only if the set E, of which ϕ is the exterior mapping function, has planar Lebesgue measure zero.

It is possible to justify our definitions of $p_n(z)$ and the α_{nk}'s when $\phi(w)$ is known only to be analytic, but not necessarily univalent, for $|w| > 1$ except for a simple pole at infinity. Pommerenke [15] showed that the validity of (4.7) for arbitrary $\langle u_n \rangle$ is *sufficient* as well as necessary for ϕ to be univalent. Sufficiency was an important part of Grunsky's original considerations [11].

The inequality (4.6) is the case of (4.7) in which $u_n = 0$, $n=1, \cdots, N-1$, and $u_N = 1$. Other inequalities stemming from (4.6) are as follows:

$$| \alpha_{nk} | \leqq (n/k)^{1/2}. \tag{4.9}$$

This is obvious from (4.6). Now let $\langle u_n \rangle$ and $\langle v_n \rangle$ be any two arbitrary sequences of complex numbers. By using the Cauchy inequality and (4.8), we find that

$$\begin{aligned} \left| \sum_{n=1}^{N} \frac{1}{n} \sum_{k=1}^{N} \alpha_{nk} u_n v_k \right| &= \left| \sum_{n=1}^{N} \left(\frac{u_n}{n^{1/2}} \sum_{k=1}^{N} \alpha_{nk} \frac{v_k}{n^{1/2}} \right) \right| \\ &\leqq \left(\sum_{n=1}^{N} \frac{|u_n|^2}{n} \right)^{1/2} \left(\sum_{n=1}^{N} \frac{1}{n} \left| \sum_{k=1}^{N} \alpha_{nk} v_k \right|^2 \right)^{1/2} \\ &\leqq \left(\sum_{n=1}^{N} \frac{|u_n|^2}{n} \right)^{1/2} \left(\sum_{n=1}^{N} \frac{|v_n|^2}{n} \right)^{1/2}. \end{aligned} \tag{4.10}$$

Grunsky's [11] classical result concerning the Faber polynomials was that this inequality with $u_n = v_n$, $n=1, 2, \cdots$, was necessary and sufficient for ϕ to be univalent.

When ∂E is an r_0-analytic simple closed curve, so ϕ is analytic and univalent for $|w| > r_0 < 1$, Pommerenke showed that sharper forms of (4.7) and (4.8) are available. Let $\phi_0(w) = \phi(r_0 w)/r_0$; this function is analytic and univalent for $|w| > 1$ except for a simple pole at infinity, and the coefficient of w in the Laurent series is again d. Let its Faber coefficients be denoted by α^*_{nk}, $n, k = 1, 2, \cdots$. From (2.10) we obtain for $|w| > 1$, $|s| > 1$:

$$\ln \frac{\phi_0(s) - \phi_0(w)}{(s-w)d} = -\sum_{n=1}^{\infty}\sum_{k=1}^{\infty} \frac{\alpha_{nk}^{*}}{n} w^{-k}s^{-n} = \ln \frac{\phi(r_0 s) - \phi(r_0 w)}{[(r_0 s) - (r_0 w)]d}$$
$$= -\sum_{n=1}^{\infty} \frac{\alpha_{nk}}{n}(r_0 w)^{-k}(r_0 s)^{-n}.$$

Therefore $\alpha_{nk} = r_0^{k+n}\alpha_{nk}^{*}$, $n,k = 1, 2, \cdots$. From (4.9) applied to $\langle \alpha_{nk}^{*}\rangle$, we obtain the inequality

$$|\alpha_{nk}| \leqq (n/k)^{1/2} r_0^{n+k}. \tag{4.3}$$

Also (4.7) and (4.10) applied to $\langle \alpha_{nk}^{*}\rangle$ with the substitution $u_n = u_n' r_0^n$, $v_n = v_n' r_0^n$, $n = 1, 2, \cdots$ yield the following theorem:

THEOREM 4.9. *Let ∂E be an r_0-analytic simple closed curve and let ϕ be analytic and univalent for $|w| > r_0 < 1$. Let $\langle u_n' \rangle$ and $\langle v_n' \rangle$ be arbitrary sequences of complex numbers. Then for each, N, $N = 1, 2, \cdots$,*

$$\sum_{k=1}^{\infty} k r_0^{-2k} \left| \sum_{n=1}^{N} \alpha_{nk} u_n' \right|^2 \leqq \sum_{n=1}^{N} n r_0^{2n} |u_n'|^2 \tag{4.11}$$

and

$$\left| \sum_{n=1}^{N} \frac{1}{n} \sum_{k=1}^{N} \alpha_{nk} u_n' v_k' \right| \leqq \left(\sum_{n=1}^{N} \frac{r_0^{2n}}{n} |u_n'|^2 \right)^{1/2} \left(\sum_{n=1}^{N} \frac{r_0^{2n}}{n} |v_n'|^2 \right)^{1/2}. \tag{4.12}$$

It would be of interest to derive inequalities of this type which are valid when ∂E is a simple closed curve with certain smoothness properties short of being analytic. At present there seems to be no information available to bridge the gap between the inequalities (4.8) and (4.10) obtained for the general case on the one hand, and the powerful inequalities (4.11) and (4.12) available for an analytic boundary ∂E.

5. The Faber series. It was observed in the Introduction that the Faber polynomials provide a simple solution to the problem of expanding a function analytic in a bounded region in a series of polynomials. Faber's original idea [9] goes along these lines: Let C be a simple closed r_0-analytic curve. Let $f(z)$ be analytic on Int C. The exterior mapping function (2.1) is analytic and univalent for $|w| > r_0 < 1$. Choose R between r_0 and 1. Then for any fixed $z \in \operatorname{Int} C_R$, by the Cauchy Integral Formula and (2.3), we have:

$$\begin{aligned} f(z) &= \frac{1}{2\pi i}\int_{C_R} \frac{f(t)dt}{t-z} = \frac{1}{2\pi i}\int_{|s|=R} \frac{f(\phi(s))\phi'(s)ds}{\phi(s) - z} \\ &= \frac{1}{2\pi i}\int_{|s|=R} f[\phi(s)]\left(\frac{1}{s} + \sum_{n=1}^{\infty} \frac{p_n(z)}{s^{n+1}}\right) ds \\ &= c_0 + \sum_{n=1}^{\infty} c_n p_n(z), \end{aligned} \tag{5.1}$$

where

$$c_n = \frac{1}{2\pi i}\int_{|s|=R} f[\phi(s)]s^{-n-1}ds, \qquad n = 0, 1, 2, \cdots. \tag{5.2}$$

The integration term-by-term implied in the last equation in (5.1) is justifiable by the fact that the power series for $\phi'(s)/[\phi(s)-z]$, with z fixed as indicated, is uniformly convergent for $|s|=R$.

The series in the last member of (5.1) is called the Faber series for f. Uniform convergence to $f(z)$ for z on any compact subset of C_R, (and therefore of Int C) is easily established by the elementary estimates (4.1).

Faber in [**9**] went on to prove by continuity considerations that if E is any bounded closed simply connected region and f is analytic on the interior points of E, then f could be approximated by a sequence of polynomials convergent throughout the interior of E, but his sequence does not identify with the partial sums of the Faber series. Apart from this result, what might be called the classical convergence theory of the Faber series, centers about two cases: (1) the boundary of E is an analytic curve C and f is analytic on the interior of E (Faber's case) and perhaps continuous throughout E; (2) E is an arbitrary compact set with simply connected complement and f is analytic at every point of E. Both these cases are adequately treated in the text-book literature; for example, see Sewell [**21**] for case (1) and Smirnov and Lebedev [**22**], Chapter 2, for case (2). Analysis in these cases is greatly facilitated by the fact that either f or the exterior mapping function ϕ can be continued analytically across the boundary.

Questions concerning the convergence of the Faber series when neither f nor ϕ can be extended analytically across the boundary are more difficult to handle, and seem to have been systematically attacked only recently. Attention has been concentrated on the situation in which E is bounded by a rectifiable simple closed curve C, and f is analytic on Int C and continuous on E. It appears that conditions beyond rectifiability are required to obtain satisfactory results. Alper and Ivanov [**3**] may have been the first to study such questions. Alper in [**2**] imposed a certain "condition j" on the boundary curve C, which is satisfied in particular whenever C is such that a tangent exists at all points and the angle between the positive real axis and the tangent to the curve, as a function of arc length, satisfies a Hölder condition. Kövari and Pommerenke [**14**] have obtained some detailed estimates and criteria for uniform convergence on E under another smoothness hypothesis on the tangent angle.

The general trend of all of these findings is that under the given hypotheses, the behavior of the Faber series on C closely reflects the behavior of the Fourier series for $f[\phi(\exp(i\theta))]$ in matters such as pointwise convergence, uniform convergence, degree of convergence, and so forth. In the remainder of this paper, we shall first develop the exact relationship between these two series and then we shall demonstrate it by establishing some easily proved convergence theorems. The methods will be much less sophisticated than those of the above references.

When E is a closed region bounded by a rectifiable simple closed curve C, and $f(z)$ is analytic on Int C, continuous on $E = C \cup \text{Int } C$, the natural reformulation of the Faber series (5.1), (5.2), is

$$f(z) \sim c_0 + \sum_{n=1}^{\infty} c_n p_n(z), \tag{5.3}$$

where now

$$\begin{aligned} c_n &= \frac{1}{2\pi i} \int_{|s|=1} f[\phi(s)] s^{-n-1} ds \\ &= \frac{1}{2\pi} \int_0^{2\pi} F(\sigma) e^{-in\sigma} d\sigma, \qquad n = 0, 1, 2, \cdots, \end{aligned}$$

and where here and in the sequel, $F(\theta) = f[\phi(\exp(i\theta))]$. The Fourier series for $F(\theta)$ in complex form is

$$F(\theta) \sim \sum_{n=-\infty}^{\infty} f_n e^{in\theta}, \tag{5.4}$$

where

$$f_n = \frac{1}{2\pi} \int_0^{2\pi} F(\sigma) e^{-in\sigma} d\sigma, \qquad n = 0, \pm 1, \pm 2, \cdots.$$

Thus $c_n = f_n$, $n = 0, 1, 2, \cdots$. This is a first link between the Faber series and the Fourier series for F.

It is worth mentioning here that the reformulation of the Faber series (5.3) can be broadly extended, both as to the generality of E and the hypothesis on f.

We proceed to complete the connection between the Faber series and the Fourier series. If C is rectifiable, then $\phi(\exp(i\theta))$ is of bounded variation as a function of θ and indeed is absolutely continuous; and it is the indefinite integral of its derivative $d\phi/d\theta$, which exists a.e. (See the argument leading to Theorem 4.4.) We introduce now the added restriction that $|d\phi/d\theta|^2$ will be Lebesgue integrable on $[0, 2\pi]$; such a rectifiable curve will be said to belong to class D^2.

LEMMA. *Let C be rectifiable and of class D^2; let f be analytic on* Int C, *continuous on $C \cup$* Int C. *Then the series $\sum_{k=1}^{\infty} f_k \alpha_{kn}$ converges, and the sum is the Fourier coefficient f_{-n}, $n = 1, 2, \cdots$.*

For the proof, we first review Parseval's formulas in Fourier series theory. Let G and H be complex-valued functions on $[0, 2\pi]$ such that G^2 and H^2 are integrable over the interval. Let $\langle C_n \rangle$ and $\langle D_n \rangle$ be their Fourier coefficients respectively. Then ([24], vol. I, p. 37)

$$\frac{1}{2\pi} \int_0^{2\pi} |G(\theta)|^2 d\theta = \sum_{n=-\infty}^{\infty} |C_n|^2, \tag{5.5}$$

$$\frac{1}{2\pi}\int_0^{2\pi} G(\theta)H(\theta)d\theta = \sum_{n=-\infty}^{\infty} C_n D_{-n}. \tag{5.6}$$

Now consider the Laurent series (2.7) for $p_n(\phi(w)) = F_n(w)$. According to Theorems 4.4 and 4.2, with the substitution $w = \exp(i\theta)$, this series becomes the Fourier series for $F_n(w)$, and moreover $F_n(w)$ is absolutely continuous.

Since $F_n(w)$ is merely a polynomial in $\phi(w)$, and $C \in D^2$, the square of the derivative $dF_n(\exp(i\theta))/d\theta$ is integrable on $[0, 2\pi]$. We substitute $w = \exp(i\theta)$ in (2.7) and formally differentiate with respect to θ term-by-term. The result is the series $(ni)\exp(ni\theta) + \sum_{k=1}^{\infty} [-ik\alpha_{nk} \exp(-ik\theta)]$. This is the Fourier series for $dF_n(\exp(i\theta))/d\theta$ ([25], vol. I, p. 40). From (5.4), (5.6), and the Grunsky Law of Symmetry we have:

$$\begin{aligned} \frac{1}{2\pi}\int_0^{2\pi} F(\theta)\frac{dF_n(\exp i\theta)}{d\theta}d\theta &= f_{-n}ni - \sum_{k=1}^{\infty} ikf_k\alpha_{nk} \\ &= f_{-n}ni - ni\sum_{k=1}^{\infty} f_k\alpha_{kn}. \end{aligned} \tag{5.7}$$

Now by the Cauchy Integral Theorem,

$$0 = \frac{1}{2\pi i}\int_C f(z)z^m dz = \frac{1}{2\pi}\int_0^{2\pi} F(\theta)[\phi(\exp(i\theta))]^m \frac{d\phi(\exp(i\theta))}{d\theta}d\theta.$$

But $dF_n[\exp(i\theta)]/d\theta$ is a linear combination of powers of $\phi(\exp(i\theta))$, multiplied by $d\phi(\exp(i\theta))/d\theta$, so the first member of (5.7) is zero. The lemma follows when ni is cancelled out of the third member.

A number of variations of the lemma are available, corresponding to other hypotheses under which Parseval's Formulas and the Cauchy Integral Theorem are valid. For example, if $f(z)$ is of bounded variation for z on C, then the hypothesis $C \in D^2$ can be dropped and the conclusion of the Lemma still holds true ([25], vol. I, p. 160). Without this hypothesis the lemma remains true if "convergence of $\sum_{k=1}^{\infty} f_k\alpha_{kn}$" is replaced by "summable by the method of arithmetic means"; that is, $(C, 1)$ summability.

Let $S_N(z)$ denote the Nth partial sum of the series (5.3), let $s_N(\theta) = S_N[\phi(\exp(i\theta))]$, and let $s_N^*(\theta) = \sum_{-N}^{N} f_n \exp(in\theta)$, which is the Nth partial sum of the Fourier series (5.4) for $F(\theta) = f[\phi(\exp(i\theta))]$. Assuming that $C \in D^2$, we have with $z = \phi(\exp(i\theta))$:

$$\begin{aligned} S_N(z) = s_N(\theta) &= c_0 + \sum_{n=1}^{N} c_n p_n(z) = f_0 + \sum_{n=1}^{N} f_n\left(e^{in\theta} + \sum_{k=1}^{\infty} \alpha_{nk}e^{-ik\theta}\right) \\ &= f_0 + \sum_{n=1}^{N} f_n e^{in\theta} + \sum_{k=1}^{N}\left(e^{-ik\theta}\sum_{n=1}^{N} f_n\alpha_{nk}\right) + \sum_{k=N+1}^{\infty}\left(e^{-ik\theta}\sum_{n=1}^{N} f_n\alpha_{nk}\right) \end{aligned}$$

$$= \sum_{n=0}^{N} f_n e^{in\theta} + \sum_{k=1}^{N}\left(e^{-ik\theta}\sum_{n=1}^{\infty} f_n\alpha_{nk}\right)$$
$$- \sum_{k=1}^{N}\left(e^{-ik\theta}\sum_{n=N+1}^{\infty} f_n\alpha_{nk}\right) + \sum_{k=N+1}^{\infty}\left(e^{-ik\theta}\sum_{n=1}^{N} f_n\alpha_{nk}\right).$$

The reversal of the order of summation after the fourth equality sign is justified by the convergence of $\sum_{k=1}^{\infty}\alpha_{nk}\exp(-ik\theta)$. The lemma justifies the fifth equality. By referring again to the lemma we obtain finally, with $z=\phi(\exp(i\theta))$,

$$S_N(z) = s_N(\theta) = \overset{*}{s}_N(\theta) + \rho_N(\theta),$$

where

$$\rho_N(\theta) = - \sum_{k=1}^{N}\left(e^{-ik\theta}\sum_{n=N+1}^{\infty} f_n\alpha_{nk}\right) + \sum_{k=N+1}^{\infty}\left(e^{-ik\theta}\sum_{n=1}^{N} f_n\alpha_{nk}\right). \tag{5.8}$$

The Riemann-Lebesgue Theorem ([25], vol. I, p. 45) states that even if $F(\theta)$ were merely integrable (instead of continuous), $\lim_{n\to\pm\infty} f_n = 0$. By Theorem 4.4, the series $\sum_{k=1}^{\infty}|\alpha_{nk}|$ is convergent. The right member of (5.8) is a trigonometric series and the series of absolute values of its coefficients is

$$\sum_{k=1}^{N}\left|\sum_{n=N+1}^{\infty} f_n\alpha_{nk}\right| + \sum_{k=N+1}^{\infty}\left|\sum_{n=1}^{N} f_n\alpha_{nk}\right|$$
$$\leqq \sum_{k=1}^{N}\left|\sum_{n=N+1}^{\infty} f_n\alpha_{nk}\right| + \sum_{k=N+1}^{\infty}\left(\sum_{n=1}^{N}|f_n\alpha_{nk}|\right)$$
$$\leqq \sum_{k=1}^{N}\left|\sum_{n=N+1}^{\infty} f_n\alpha_{nk}\right| + \left(\max_j |f_j|\right)\sum_{n=1}^{N}\left(\sum_{k=N+1}^{\infty}|\alpha_{nk}|\right).$$

It follows that the trigonometric series in (5.8) converges absolutely and therefore uniformly to $\rho_N(\theta)$ for all θ. We have:

THEOREM 5.1. *With $C\in D^2$ and f continuous on $C\cup\mathrm{Int}\, C$, analytic on* Int C, *the difference $\rho_N(\theta)=s_N(\theta)-\overset{*}{s}_N(\theta)$ between the Nth partial sum of the Faber series for f with $z=\phi(\exp i\theta)$ and the Nth partial sum of the Fourier series for $f[\phi(\exp(i\theta))]$ is the absolutely and uniformly convergent Fourier series appearing in* (5.8).

The theorem has the limitation that if it is to be used to obtain Faber convergence theorems from Fourier theorems, certain global estimates of the α_{nk} seem to be needed, and these can be difficult to interpret geometrically as conditions on C.

The case in which C is an r_0-analytic curve presents no problems, because then it follows from the estimates (4.3) that there exists a constant $M>0$ depending on f and C such that $|\rho_N(\theta)| = Mr_0^N$ for all θ, where $0\leqq r_0<1$.

We conclude by exhibiting two relatively simple convergence theorems which can be proved by means of Theorem 5.1.

THEOREM 5.2. *Let $C \in D^2$ be such that a constant $M > 0$ exists with the property that*

$$\int_0^{2\pi}\int_0^{2\pi}\left|\frac{\partial}{\partial s}\ln\frac{\phi(s)-\phi(w)}{s-w}\right|^2 d\sigma d\theta \leqq M, \quad s = Re^{i\sigma}, \quad w = Re^{i\theta}$$

for all $R > 1$. Let $f(z)$ be analytic on Int C, *continuous on* $C \cup$ Int C, *and let $S_N(z)$ denote the Nth partial sum of its Faber series. Then*

$$\operatorname*{Lim}_{N\to\infty}\int_C |S_N(z) - f(z)|\,|dz| = 0$$

and $\operatorname{Lim}_{N\to\infty} S_N(z) = f(z)$ *uniformly on any compact subset of* Int C.

The second limit follows directly from the first by the Cauchy Integral Formula:

$$S_N(z) - f(z) = \frac{1}{2\pi i}\int_C \frac{S_N(t) - f(t)}{t - z}\,dt.$$

As for the first limit, with the notation introduced for Theorem 5.1,

$$\begin{aligned}
\int_C |S_N(z) - f(z)|\,|dz| &= \int_0^{2\pi} |s_N(\theta) - F(\theta)|\left|\frac{d\phi}{d\theta}\right| d\theta \\
&\leqq \int_0^{2\pi} |s_N(\theta) - \overset{*}{s}_N(\theta)|\left|\frac{d\phi}{d\theta}\right| d\theta + \int_0^{2\pi} |\overset{*}{s}_N(\theta) - F(\theta)|\left|\frac{d\phi}{d\theta}\right| d\theta \\
&\leqq \left[\int_0^{2\pi} |s_N(\theta) - \overset{*}{s}_N(\theta)|^2 d\theta \int_0^{2\pi}\left|\frac{d\phi}{d\theta}\right| d\theta\right]^{1/2} \\
&\quad + \left[\int_0^{2\pi} |\overset{*}{s}_N(\theta) - F(\theta)|^2 d\theta \int_0^{2\pi}\left|\frac{d\phi}{d\theta}\right|^2 d\theta\right]^{1/2}.
\end{aligned}$$

The Schwarz-Buniakowsky inequality was used twice at the last step. It is clear that the desired result can be achieved by proving that the first integral in each square brackets tends to zero as $N \to \infty$. By Parseval's Formula (5.5), we have

$$\frac{1}{2\pi}\int_0^{2\pi} |\overset{*}{s}_N(\theta) - F(\theta)|^2 d\theta = \left(\sum_{n=-N-1}^{-\infty} + \sum_{n=N+1}^{\infty}\right)|f_n|^2$$

which must approach zero as $N \to \infty$, by Parseval's Formula applied to F above. Again using Parseval's Formula and thereafter the Cauchy Inequality,

$$\begin{aligned}
\frac{1}{2\pi}\int_0^{2\pi} |s_N(\theta) - \overset{*}{s}_N(\theta)|^2 d\theta &= \frac{1}{2\pi}\int_0^{2\pi} |\rho_N(\theta)|\,d\theta \\
&= \sum_{k=1}^{N}\left|\sum_{n=N+1}^{\infty} f_n\alpha_{nk}\right|^2 + \sum_{k=N+1}^{\infty}\left|\sum_{n=1}^{N} f_n\alpha_{nk}\right|^2
\end{aligned} \tag{5.9}$$

$$
\begin{aligned}
(5.9) \qquad &\leqq \sum_{n=N+1}^{\infty} |f_n|^2 \sum_{k=1}^{N} \sum_{n=N+1}^{\infty} |\alpha_{nk}|^2 + \sum_{n=1}^{N} |f_n|^2 \sum_{k=N+1}^{\infty} \sum_{n=1}^{N} |\alpha_{nk}| \\
&\leqq \sum_{n=N+1}^{\infty} |f_n|^2 \sum_{k=1}^{\infty} \sum_{n=1}^{\infty} |\alpha_{nk}|^2 + \sum_{n=1}^{\infty} |f_n|^2 \sum_{k=N+1}^{\infty} \sum_{n=1}^{\infty} |\alpha_{nk}| .
\end{aligned}
$$

According to Theorem 4.5, the present hypotheses imply that $\sum_{k=1}^{\infty} \sum_{n=1}^{\infty} |\alpha_{nk}|^2 < \infty$. From this, and Parseval's Formula applied to F, it follows that each term in the last member of (5.9) tends to zero as $N \to \infty$.

Theorem 4.6 provides a sufficient condition in terms of derivatives of ϕ for the hypothesis on C of Theorem 5.2 to be valid.

A more naive result is as follows:

THEOREM 5.3. *Let $C \in D^2$ be such that $\sum_{n=1}^{\infty} (\sum_{k=1}^{\infty} |\alpha_{nk}|)^2 < \infty$. Let $f(z)$ be analytic on* Int C, *continuous on* $C \cup \text{Int } C$. *Then $S_n(z)$ for $z = \phi(\exp(i\theta))$ and $s_N^*(\theta)$ are pointwise equiconvergent, and if $s_N^*(\theta)$ converges uniformly on a set $T \subset [0, 2\pi]$, then $S_n(z)$ converges uniformly on $\phi(T)$.*

The proof consists in rearranging (5.8) in the form

$$\rho_N(\theta) = - \sum_{n=N+1}^{\infty} \left(f_n \sum_{k=1}^{N} \alpha_{nk} e^{-ik\theta} \right) + \sum_{n=1}^{N} \left(f_n \sum_{k=N+1}^{\infty} \alpha_{nk} e^{-ik\theta} \right),$$

which is justifiable, and then applying the Cauchy inequality and Parseval's Formula for F.

In conclusion, it should be mentioned that Rosenbloom and Warschawski in [18] announced a theorem similar to Theorem 5.2 above, but for a class of modified Faber polynomials (apparently first introduced by Szegö [24]) which are respectively the principal part, not of $[\phi^{-1}(z)]^n$, but of $[\phi^{-1}(z)]^n [d\phi^{-1}/dz]^{1/2}$. Rosenbloom and Warschawski replace our continuity conditions on the approximated function f by a condition which insures that f shall have suitably integrable boundary values on C. A similar generalization of Theorem 5.2 is possible.

The preparation of this paper was supported by the Air Force Office of Scientific Research, Office of Aerospace Research, United States Air Force, under AFOSR Grant 68-1525.

References

1. L. V. Ahlfors, Complex Analysis, sec. ed., McGraw-Hill, New York, 1966.

2. S. Ya Al'per, On uniform approximations of functions of a complex variable in a closed region, Izv. Akad. Nauk SSSR. Ser. Math., 19 (1955) 423–444.

3. S. Ya Al'per and V. V. Ivanov, On the approximation of functions by partial sums of a series of Faber polynomials, Doklady Akad. Nauk SSSR (N.S.), 90 (1953) 325–328.

4. Lennart Carleson, On convergence and growth of partial sums of Fourier series, Acta Math., 116 (1966) 135–157.

5. Z. Charzynski and M. Schiffer, A new proof of the Bieberbach conjecture for the fourth coefficient, Arch. Rational Mech. Anal., 5 (1960) 187–193.

6. J. H. Curtiss, Riemann sums and the fundamental polynomials of Lagrange interpolation, Duke Math. J., 8 (1941) 525–532.

7. J. H. Curtiss, Harmonic interpolation in Fejér points with Faber polynomials as a basis, Math. Z., 86 (1964) 75–92.

8. ———, Transfinite diameter and harmonic polynomial interpolation, J. Analyse Math., 22 (1969) 371–389.

9. Georg Faber, Über polynomische Entwicklungen, Math. Ann., 57 (1903) 398–408.

10. ———, Über Tschebyscheffsche Polynome, J. Reine Angew. Math., 150 (1920) 79–106.

11. Helmut Grunsky, Koeffizientenbedingungen für schlicht abbildende meromorphe Funktionen, Math. Z., 45 (1939) 29–61.

12. Einar Hille, Analytic Function Theory, Ginn, Boston, 1962.

13. J. A. Jenkins, Some area theorems and a special coefficient theorem, Ill. J. Math., 8 (1964) 80–99.

14. T. Kövari and Ch. Pommerenke, On Faber polynomials and Faber expansions, Math. Z., 99 (1967) 193–206.

15. Ch. Pommerenke, Über die Faberschen Polynome schlichter Funktionen, Mathematische Z., 85 (1964) 197–208.

16. ———, Konform Abbildung und Fekete-Punkte, Math. Z., 89 (1965) 422–438.

17. ———, Über die Verteilung der Fekete-Punkte, Math. Ann., 168 (1967) 111–127.

18. P. C. Rosenbloom and S. E. Warschawski, Approximation by polynomials, Lectures on Functions of a Complex Variable, ed. Wilfred Kaplan et al., the University of Michigan Press, 1955.

19. Menahem Schiffer, Faber polynomials in the theory of univalent functions, Bull. Am. Math. Soc., 54 (1948) 503–517.

20. I. Schur, On Faber Polynomials, Amer. J. Math., 67 (1945) 33–41.

21. W. E. Sewell, Degree of Approximation by Polynomials in the Complex Domain, Princeton University Press, 1942.

22. V. I. Smirnov and N. A. Lebedev, Functions of a Complex Variable-Constructive Theory, the M.I.T. Press, 1968.

23. P. K. Suetin, The basic properties of Faber polynomials (Russian), Uspehi Mat. Nauk, No. 4 (118), 19 (1964) 125–154.

24. Gabor Szegö, Orthogonal Polynomials, Amer. Math. Soc., 1939.

25. Antoni Zygmund, Trigonometric Series, Cambridge University Press, 1959.

RINGS OF FRACTIONS

P. M. COHN, Bedford College, London

Introduction. A well-known (and easily proved) theorem states that each integral domain can be embedded in a field [37]. This was generalized to certain noncommutative rings in a brilliant paper by Ore [33] in 1930; his results are essentially as simple as in the commutative case, and the proofs, though longer, are no harder. Beyond this, very little is known, so little that it can be set down in quite a brief article. I thought this was worth doing because some interesting

Professor Cohn did his Cambridge Ph.D. under Philip Hall, and he has held positions at the Univ. de Nancy, Manchester University, Queen Mary College London, and (presently) Bedford College, London. He has spent leaves-of-absence at Yale Univ., Univ. of Chicago, and Rutgers. He has published extensively in universal algebra and in many branches of algebra; in 1965–67 he was the Secretary of the London Mathematical Society. His books are *Lie Groups* (Cambridge Univ. Press 1957), *Linear Equations* (Routledge and Kegan Paul 1958), *Solid Geometry* (Routledge and Kegan Paul 1961), and *Universal Algebra* (Harper and Row 1965). *Editor.*

problems on the embedding of rings in skew fields remain, and because it gives me the chance to mention some recent work which makes it seem that these embedding problems are not quite as hard as they appear at first sight.

The central problem, finding *fields* of fractions, is really part of the problem of constructing *rings* of fractions (i.e., inverting certain elements), which also has other important applications. We shall therefore arrange the discussion so as to include this more general case.

Conventions. Every ring has a unit-element, denoted by 1, which is preserved by homomorphisms, inherited by subrings and acts as the identity operator on modules. The same conventions apply to semigroups. Of course a ring may very well consist of 0 alone; this is the case precisely when $1=0$. We use 0 to denote both the zero element and the set consisting of the zero element; the context will always make clear which is intended.

In any ring R the set of nonzero elements is denoted by R^*. A ring R, not necessarily commutative, such that R^* is a group under multiplication will be called a *field*. In the current literature this is often called a *skew field* or *division ring*; we shall occasionally use the prefix 'skew' for emphasis. A ring R such that R^* is a semigroup under multiplication is said to be *entire*, and a commutative entire ring is called an *integral domain*. Note that in a field $1 \neq 0$; the same is true in entire rings, by our convention about semigroups.

An element u in a ring is *invertible* or a *unit* if it has an *inverse* u^{-1} satisfying $uu^{-1}=u^{-1}u=1$; of course the inverse is unique if it exists at all. In an entire ring, if $uv=1$, then $u(vu-1)=(uv-1)u=0$, hence $vu=1$ and v is the inverse of u. Thus all one-sided inverses are two-sided in this case. An element u is called a *zero-divisor* if $u \neq 0$ and if for some $v \neq 0$, either $uv=0$ or $vu=0$. A *nonzero-divisor* is a nonzero element which is not a zero-divisor. Thus by our convention 0 is neither a zero-divisor nor a nonzero-divisor.

If R is any ring, a *field of fractions* of R is a field containing R as a subring and generated, as a field, by R.

Outline. In Section 1 we review the commutative case; Section 2 introduces the obvious but rather useful notion of a 'universal S-inverting ring' and also gives Malcev's example of an entire ring not embeddable in a field. The Ore construction occupies Section 3, with applications in Section 4, including the theorems of Goldie and Posner. The remaining sections describe methods of constructing fields of fractions which go beyond Ore's theorem. In Section 5 we examine generalized inverses, and the relation to the Johnson-Utumi 'ring of quotients'; this turns out to be not very close in the case of chief interest to us, that of entire rings. The topological methods in Section 6 essentially generalize the notion of a decimal fraction. The final Section 7 reports briefly on the author's recent method of embedding rings in fields by inverting matrices rather than elements.

1. The commutative case. Let R be a commutative ring. The need for

fractions arises when we try to enlarge R so as to ensure that equations of the form

$$xb = a \tag{1}$$

can be solved. At this stage our instinct should tell us to beware of the case $b=0$, but we shall leave this question aside for the moment. If we denote the solution of (1) by ab^{-1} or a/b, we see immediately that $a/b = ac/bc$, or more generally,

$$a/b = a'/b' \qquad \text{if and only if } ab' = ba', \tag{2}$$

under suitable restrictions to exclude division by 0. Further, if the solutions are to form a ring containing R, they must add and multiply according to the rules

$$\frac{a}{b} + \frac{a'}{b'} = \frac{ab' + a'b}{bb'}, \qquad \frac{a}{b}\cdot\frac{a'}{b'} = \frac{aa'}{bb'}. \tag{3}$$

We learn at an early stage of our algebra course that if R is an integral domain, then we can find a field containing R as subring by taking all fractions a/b with $b \neq 0$ and combining them according to the rules (3), bearing in mind the cancellation rule (2).

To generalize this construction we observe that to form the new denominators in (3), we need only *multiply* the denominators b and b' together, not add them; this suggests taking a subsemigroup S of R as our stock of denominators. Now we shall in general no longer obtain a field; we may not even get a ring containing R as subring, but by following essentially the same construction we get a ring R_S say, with a homomorphism

$$\lambda : R \to R_S$$

which maps the elements of S to invertible elements, and it will be a simple matter to find out when λ is injective.

Thus we are given a subsemigroup S of a commutative ring R and we define a relation on the product set $R \times S$ by the rule:

$$(a, s) \sim (a', s') \text{ if and only if } as't = a'st \text{ for some } t \in S. \tag{4}$$

This reduces to (2) when S consists of nonzero-divisors; in general the form (4) is necessary to make sure that '$\sim$' is really an equivalence relation. Let us only check transitivity (reflexivity and symmetry are obvious): If $(a, s) \sim (a', s')$ and $(a', s') \sim (a'', s'')$, then $as't = a'st$ and $a's''t' = a''s't'$ for some $t, t' \in S$; hence

$$as'' \cdot s'tt' = a'ss''tt' = a''s'stt' = a''s \cdot s'tt'.$$

Since $s'tt' \in S$, this shows that $(a, s) \sim (a'', s'')$ and the transitivity is proved. We denote the equivalence class containing (a, s) by a/s and define addition and multiplication by the formulae (3); of course we must verify that the definitions really only depend on the classes of a/s, a'/s' and not on the representatives

of these classes used in the formulae. This is a routine verification, as is the proof that the set of classes a/s with these operations forms a ring, denoted by R_S say, with zero $0/1$ and unit-element $1/1$. The mapping

$$\lambda : a \mapsto a/1 \tag{5}$$

of R into R_S is clearly a homomorphism; it maps each element of S to an invertible element of R_S, because $(s/1)(1/s) = s/s = 1/1$.

Let us say that a homomorphism $f:R \to R'$ is *S-inverting* if each element of S is mapped by f to an invertible element of R'. For example the mapping (5) is S-inverting, but we can say more than this. Let $f:R \to R'$ be any S-inverting homomorphism and define a mapping $f_1: R \times S \to R'$ by the rule

$$(a, s)^{f_1} = a^f(s^f)^{-1}.$$

This makes sense because f is S-inverting, by hypothesis. We observe now that f_1 takes the same value on pairs that are equivalent according to (4): If $as't = a'st$, then $a^f s'^f t^f = a'^f s^f t^f$ and hence $a^f(s^f)^{-1} = a'^f(s'^f)^{-1}$. This means that we obtain a well-defined mapping f' of R_S into R' by putting $(a/s)^{f'} = (a, s)^{f_1}$. This mapping f' has the property that for any $a \in R$,

$$(a/1)^{f'} = a^f, \tag{6}$$

an equation which may also be expressed by saying that the accompanying diagram commutes, i.e., $\lambda f' = f$. Moreover, f' is uniquely determined by (6), because

$$\begin{tikzcd} R \arrow[r, "\lambda"] \arrow[dr, "f"'] & R_S \arrow[d, "f'"] \\ & R' \end{tikzcd}$$

that equation determines its value on the elements $a/1$, and its value on $1/s$ must be the inverse of its value on $s/1$. Thus the mapping

$$\lambda : R \to R_S \tag{7}$$

is not just an S-inverting homomorphism, but the most general such homomorphism, in the sense that each S-inverting homomorphism can be obtained by taking a uniquely determined homomorphism from R_S. This property is expressed by saying that (7) is the *universal S-inverting homomorphism*, and R_S itself is called the *universal S-inverting ring*. This universal property in effect determines R_S up to isomorphism.

We shall wish to know when λ is injective; more generally, let us determine the kernel of λ. Clearly $a^\lambda = 0$ if and only if $a/1 = 0/1$, and by definition this means that $at = 0$ for some $t \in S$. We can now sum up our results:

THEOREM 1.1. *Let R be a commutative ring and S a subsemigroup of R. Then*

there is a universal S-inverting homomorphism $\lambda: R \to R_S$; *the elements of* R_S *can be written as fractions* a/s $(a \in R, s \in S)$, *where* $a/s = a'/s'$ *if and only if* $as't = a'st$ *for some* $t \in S$. *Further,*

$$\ker \lambda = \{a \in R \mid at = 0 \text{ for some } t \in S\}.$$

REMARKS. 1. Note that R_S is again commutative.

2. The ring R_S reduces to 0 precisely when $0 \in S$; it is to avoid this trivial case that one usually assumes $0 \notin S$.

3. The mapping λ is injective if and only if S contains only nonzero-divisors. In that case R_S is called a *ring of fractions* of R; the largest such ring is obtained by taking S to be the set of all nonzero-divisors, a set which is always a subsemigroup. The ring obtained by inverting all nonzero-divisors is called the *total ring of fractions* of R. In case R is an integral domain, this is the universal R^*-inverting ring of R, which is of course the field of fractions of R.

An important special case of the theorem is obtained by taking S to be the complement of a prime ideal $\mathfrak{p}$ in R (a *prime* ideal is an ideal of R whose complement is a semigroup under multiplication; note that this does not allow R as prime ideal). In that case one often writes $R_{\mathfrak{p}}$ instead of R_S, somewhat inconsistently, but without risk of confusion, because S never contains 0, whereas $\mathfrak{p}$ always does.

The problem of constructing fractions already arises in a semigroup, and should really be considered in that setting. We have nevertheless treated the case of rings first, on account of its importance; it also happens to coincide with the historical order of its development [16]. In any case we can easily extract the answer for semigroups from our conclusion:

THEOREM 1.2. *Let* M *be a commutative semigroup and* S *a subsemigroup of* M. *Then there is a universal S-inverting homomorphism*

$$\lambda: M \to M_S,$$

the elements of M_S *can be written as fractions* a/s $(a \in M, s \in S)$ *as in the ring case, and* a, a' *have the same image under* λ *if and only if* $at = a't$ *for some* $t \in S$.

2. Some observations on the general case. Let us return to our basic problem, which is to construct a field of fractions for a given ring, when possible. If a ring R is to be embedded in a field, then whether commutative or not, R must be entire. But this necessary condition, which in the commutative case was sufficient, in general is no longer so. The first example of an entire ring not embeddable in a field was given by Malcev [28]. He takes the ring R generated by eight elements a, b, c, d, x, y, u, v, with defining relations

$$ax = by, \qquad cx = dy, \qquad cu = dv. \tag{1}$$

To show that this ring is entire, one uses a normal form for its elements. In outline the argument goes as follows. Each element of R can, by use of (1), be expressed as a noncommutative polynomial in the given generators, in which

there are no occurrences of by, dy, dv (the right-hand sides of the equations (1)), and such an expression is unique. The verification that the product of nonzero elements is nonzero is fairly straightforward, though care is needed to ensure that all possibilities are considered at each stage. The normal form also shows that $au \neq bv$, but if R were embeddable in a field, or even in a ring in which a, c, y, v have inverses, we could deduce from (1) that $a^{-1}b = xy^{-1}$, $xy^{-1} = c^{-1}d$, $c^{-1}d = uv^{-1}$, hence $a^{-1}b = uv^{-1}$, and so

$$au = bv, \tag{2}$$

which is a contradiction.

Malcev obtained his example in the course of studying conditions under which a semigroup is embeddable in a group [**29**]. In fact he was able to write down an infinite series of 'quasi-identities', i.e., conditions of the form

$$A_1, \cdots, A_n \text{ imply } B,$$

(where $A_1, \cdots, A_n, B$ are equations in a semigroup) which he proved necessary and sufficient for a semigroup to be embeddable in a group. The simplest of these quasi-identities are left and right cancellation: '$xy = xz$ implies $y = z$' and '$xz = yz$ implies $x = y$.' The next condition is of the form 'the equations (1) imply (2)', and the example just given shows it to be independent of cancellation (more generally, the infinite set of quasi-identities cannot be replaced by any finite subset). A detailed account of Malcev's Theorem can be found in [**11**]; it seems likely that any corresponding criterion for the embeddability of rings in skew fields is rather more complicated. But it follows from results in general algebra that the embeddability of a nonzero ring in a field can be expressed by a (possibly infinite) set of quasi-identities ([**11**], p. 235).

Recently, in [**41**], A. A. Klein has found an infinite set of quasi-identities which are necessary for embeddability in a field and which he conjectures to be sufficient.* They express that R is entire and for all n, each nilpotent $n \times n$ matrix C satisfies $C^n = 0$.

Malcev has asked whether rings exist whose nonzero elements are embeddable in groups, but which are not embeddable in fields. Such rings were found simultaneously and independently by three people in 1966 [**4, 5, 24**].

Let us now take the general situation and see whether anything found in the commutative case can be used here. If R is a ring and S any subset (not necessarily a subsemigroup), we can define an S-inverting homomorphism as before. Given an S-inverting homomorphism $f: R \to R'$, let $\bar{S}$ be the subset of R whose elements are mapped into invertible elements of R'. Clearly $\bar{S} \supseteq S$, but equality need not hold; in particular $\bar{S}$ contains all invertible elements of R, and if $u, v \in \bar{S}$, then $uv \in \bar{S}$, because $(uv)^f = u^f v^f$, and the latter is invertible when u^f, v^f are. This shows that $\bar{S}$ is always a subsemigroup, so nothing is lost by taking the set to be inverted as a semigroup.

* Added in proof (March 29, 1971): A counterexample to this Conjecture has just been found by G. M. Bergman.

We can again construct a universal S-inverting ring R_S: simply take a presentation of R (by generators and defining relations) and for each $s \in S$ adjoin a new generator s' and extra relations

$$ss' = s's = 1. \tag{3}$$

The ring R_S so obtained may no longer contain R as subring, e.g., if we apply this construction to Malcev's example, with $S = \{a, c, y, v\}$, we get a collapse because then $au = bv$. But we always have a natural homomorphism $R \to R_S$ which is S-inverting, and in fact this is the universal S-inverting homomorphism, because the relations holding in R_S, namely (3) together with the relations of R itself, must hold in any image of R under an S-inverting homomorphism. (This is essentially an application of Dyck's Theorem, see, e.g., [11], p. 183.) In this way we obtain the following result:

THEOREM 2.1. *Let R be any ring and S any subset of R. Then there is a universal S-inverting homomorphism $\lambda: R \to R_S$, where R_S is unique up to isomorphism. Moreover, λ is injective if and only if R can be embedded in a ring in which all elements of S have inverses.*

The assertion is quite general and, because of its very generality, rather easy to prove. It is also not hard to see that the correspondence $(R, S) \mapsto R_S$ is a *functor* (from the category of pairs R, S to the category of rings). This means that to each homomorphism of rings $f: R \to R'$ such that $S^f \subseteq S'$ for subsets S, S' of R, R' respectively, there corresponds a homomorphism $\bar{f}: R_S \to R'_{S'}$ such that $\overline{fg} = \bar{f}\bar{g}$ and the identity mapping on R corresponds to the identity on $R_S: \bar{1} = 1$. Moreover, the natural mappings $\lambda: R \to R_S$ and $\lambda': R' \to R'_{S'}$ have the property of making the following diagram commute:

$$\begin{tikzcd} R \arrow[r, "f"] \arrow[d, "\lambda"'] & R' \arrow[d, "\lambda'"] \\ R_S \arrow[r, "\bar{f}"] & R'_{S'} \end{tikzcd}$$

This is expressed by saying that λ is a *natural transformation* (for details cf. e.g., [27]).

At first sight Th. 2.1 looks deceptively like Th. 1.1, but it has the serious drawback that no normal form for the elements of R_S is given. This makes it hard to decide when λ is injective; also we cannot be sure, even when R is embeddable in a field, that R_{R^*} will be the whole field of fractions. The trouble is that after adjoining inverses of all the nonzero elements to R, there may still be elements without inverses, e.g., elements of the form $ab^{-1}c + de^{-1}f$. So the process of adjoining inverses may have to be repeated, perhaps infinitely often. The following observation is sometimes useful:

THEOREM 2.2. *Let R be any ring. If there is an R^*-inverting homomorphism*

of R into a field K, then R is embeddable in a field.

For by hypothesis, we have a homomorphism

$$f: R \to K, \tag{4}$$

and since any nonzero element of R maps to an invertible element of K, it cannot map to 0, i.e., (4) is injective.

The necessity of having to repeat the process of adjoining inverses does not arise for semigroups: If a semigroup M is embeddable in a group G, then the subsemigroup of G generated by the elements of M and their inverses already forms a group. Neither does the problem arise in the special case treated by Ore, to which we now turn.

3. Ore's Construction. In an attempt to carry over the results of Section 1 to the noncommutative case, let us examine the situation where every element of the universal S-inverting ring R_S can be written in the form of a fraction a/s. If this is to be possible, we must be able to express $(1/s)(a/1)$ in this form, say

$$(1/s)(a/1) = a'/s'. \tag{1}$$

Multiplying both sides by $s/1$ on the left and by $s'/1$ on the right, we get

$$as'/1 = sa'/1. \tag{2}$$

This gives us a clue to the extra condition required now.

THEOREM 3.1. *Let R be any ring, S a subsemigroup, and assume further that*

(i) *for any $a \in R$ and $s \in S$, $aS \cap sR \neq \emptyset$,*

(ii) *for any $a \in R$ and $s \in S$, if $sa = 0$, then $at = 0$ for some $t \in S$.*

Then the elements of the universal S-inverting ring R_S can be constructed as fractions a/s ($a \in R$, $s \in S$), where

$$a/s = a'/s' \Rightarrow au = a'u', \qquad su = s'u' \in S \text{ for some } u, u' \in R. \tag{3}$$

The kernel of the natural mapping is then

$$\ker \lambda = \{a \in R \mid at = 0 \text{ for some } t \in S\}.$$

Of course here we must distinguish carefully between as^{-1} and $s^{-1}a$; the expression a/s corresponds to the former.

A subsemigroup S of R satisfying the conditions (i), (ii) of this theorem will be called a *right denominator set* in R. The proof of this result is largely an exercise in patience, and the reader is recommended to verify at least some of the steps. The basic observation is that any two fractions can be brought to a common denominator in S, using (i), and they represent the same element of R_S if and only if, over a suitable common denominator, their numerators are equal (cf. (2)). Addition of fractions over the same denominator is straightforward, and the multiplication of fractions is based on the rule (1).

REMARKS: 1. Again $R_S = 0$ if and only if $0 \in S$; one usually excludes this case.

2. There is a left-right analogue of the theorem, obtained by switching sides; it shows how to form *left* fractions, starting from a *left* denominator set.

3. There is a corresponding theorem for constructing fractions in a semigroup. More generally, the construction can be performed in any category (cf. [**20**], p. 28).

4. Any subsemigroup S consisting of invertible elements of R is a right (and left) denominator set, and the universal S-inverting homomorphism is then an isomorphism.

5. Any *central* subsemigroup S (i.e., satisfying $as=sa$ for all $a\in R$, $s\in S$) is a right (and left) denominator set.

6. The universal S-inverting mapping is injective if and only if S contains only nonzero-divisors. In that case condition (ii) becomes superfluous. This case is sufficiently important to be stated separately.

COROLLARY 1. *Let R be a ring and S a subsemigroup of R consisting of nonzerodivisors, such that $aS\cap sR\neq\emptyset$ for any $a\in R$, $s\in S$. Then the universal S-inverting homomorphism is injective.*

When S is as in Corollary 1, R_S is again called a *ring of fractions*, *total* in case S consists of all nonzero-divisors. But this time we cannot be sure that there is a total ring of fractions, because the set of all nonzero-divisors need not satisfy the hypothesis of Corollary 1.

If R is entire and R^* is a right denominator set, we get the case originally treated by Ore (cf. [**33**]; the generalizations were given by Asano [**3**] and others).

COROLLARY 2. *Let R be an entire ring such that*

$$aR\cap bR\neq 0 \qquad \text{for any } a,\, b\in R^*. \tag{4}$$

Then R can be embedded in a skew field K, whose elements have the form of fractions a/b ($a\in R$, $b\in R^*$).

Condition (4) is called the *right Ore condition*, and an entire ring satisfying (4) is called a *right Ore ring*. We observe that (4) is necessary as well as sufficient for the conclusion to hold. For if an entire ring R can be embedded in a field K in such a way that each element of K has the form ab^{-1} (a, $b\in R$), then in particular, for any a, $b\in R^*$ we can find a', $b'\in R^*$ such that $b^{-1}a=a'b'^{-1}$; hence $ab'=ba'\neq 0$.

4. Applications of Ore's Construction. An important class of Ore rings (which Ore himself had studied and clearly had in mind when making his construction, cf. [**34**]) is formed by the *skew polynomial rings*. Given any field K, passing to the polynomial ring $K[x]$ in an indeterminate x is a familiar construction, which still works even when K is skew. In that case it is usual to require x to be *central*, i.e., to commute with the field elements. But we often need a more general case: we still assume that each element of our ring can be written as a polynomial in just one way:

$$f = a_0 + xa_1 + \cdots + x^n a_n \qquad (a_i \in K); \tag{1}$$

we no longer assume that x is central, but instead that for each $a \in K$ there exist $\bar{a}$, $a' \in K$ satisfying

$$ax = x\bar{a} + a'. \tag{2}$$

It is not too hard to show that the mapping $\alpha: a \mapsto \bar{a}$ is an endomorphism of K, and that $\delta: a \mapsto a'$ is an additive mapping satisfying

$$(ab)' = a'\bar{b} + ab'. \tag{3}$$

Any additive mapping δ of a field into itself, satisfying (3) (for some endomorphism $\alpha: a \mapsto \bar{a}$) is called an *$\alpha$-derivation*. E.g., on the field of rational functions $F(t)$ over some commutative field F, the usual derivative $f' = df/dt$ defines a 1-derivation (associated with the identity automorphism of $F(t)$).

Conversely, given any endomorphism α of a field K and any α-derivation δ, we can define a multiplication on the set of polynomials (1) by using the commutation rule (2). This leads to a ring denoted by $K[x; \alpha, \delta]$ and called a *skew polynomial ring*. This ring is entire and is in fact a right Ore ring; for the proof we can use the Euclidean algorithm, as in ordinary polynomial rings, but we must take care here to perform all divisions on the right. It follows that we can form the field of fractions, denoted by $K(x; \alpha, \delta)$. Of course for $\alpha = 1$, $\delta = 0$ the skew polynomial ring reduces to the ordinary polynomial ring $K[x]$ with a central indeterminate and its field of fractions $K(x)$.

It is important to note that the construction just given is unsymmetric, and $K[x; \alpha, \delta]$ will not in general be a *left* Ore ring. The condition for it to be one is that α should be an automorphism, for this is the condition which enables us to rewrite (2) as a commutator formula in the other direction:

$$xa = a_1 x + a_2. \tag{4}$$

Explicitly, if $ax = xa^\alpha + a^\delta$ and $\alpha^{-1} = \beta$, then $a^\beta x = xa + a^{\beta\delta}$, hence (4) holds with $a_1 = a^\beta$, $a_2 = -a^{\beta\delta}$. Conversely, if α is not an automorphism, take a in K but not in the image under α. Then it is easily checked that x and xa have no common left multiple other than 0; so $K[x; \alpha, \delta]$ is a *left* Ore ring precisely when α is an automorphism.

We observe that the class of right Ore rings is closed under forming polynomial rings [**15**].

THEOREM 4.1. *If R is a right Ore ring, then so is the ring $R[x]$ of polynomials in a central indeterminate.*

Proof. Since R is right Ore, it has a field of fractions, K say, and $R[x]$ can be embedded in the field $K(x)$ of rational functions in x. Each element of $K[x]$ has the form fa^{-1}, where $f \in R[x]$ and $a \in R^*$ is a common denominator for the coefficients. Hence each element w of $K(x)$ can be written $w = fa^{-1}(gb^{-1})^{-1} = fa^{-1}bg^{-1}$, where f, $g \in R[x]$ and a, $b \in R^*$. Since R is right Ore, it contains a', b' such that $ab' = ba' \neq 0$; hence $a^{-1}b = b'a'^{-1}$, so $w = fb'a'^{-1}g^{-1} = fb'(ga')^{-1}$. This shows that

each element of $K(x)$ is a fraction of elements of $R[x]$; therefore the latter is a right Ore ring, as asserted.

It is a remarkable fact, first observed by Goldie [18], that every right Noetherian entire ring is a right Ore ring. (A ring is *right Noetherian* if all its right ideals are finitely generated.) For if R is entire and right Noetherian, take $x, y\neq 0$, and consider the right ideal generated by the elements x^ny ($n=0, 1, \cdots$). This must be finitely generated, say $x^ny=ya_0+xya_1+\cdots+x^{n-1}ya_{n-1}$. If $a_0=0$, we can cancel a power of x from the left, so we may assume that $a_0\neq 0$, and then

$$x(x^{n-1}y - ya_1 - \cdots - x^{n-2}ya_{n-1}) = ya_0 \neq 0.$$

The result is sometimes called the "little Goldie theorem":

THEOREM 4.2. *Every right Noetherian entire ring is a right Ore ring.*

Similarly the total ring of fractions introduced earlier plays a role in the "big Goldie theorem." To state it we recall that a ring R is said to be *prime* if the product of nonzero ideals in R is nonzero and *semiprime* if the square of each nonzero ideal is nonzero. A ring is *right Artinian* if its right ideals satisfy the descending chain condition; this is a very much stronger condition than being right Noetherian, and much more is known about the Artinian rings (cf. e.g., [26]). Goldie's theorem provides a connection between the two; in one form it states:

If a right Noetherian ring R is prime (*respectively semiprime*), *then R has a total ring of fractions which is right Artinian and simple* (*respectively semisimple*).

For a brief proof see [19].

Let us return to Th.4.2 and consider an entire ring R which is not right Ore. For simplicity we take R to be an algebra over a commutative field F. By hypothesis we can find $x, y\in R^*$ such that $xR\cap yR=0$. It follows that there is no polynomial in x and y (treated as noncommuting variables) which is zero, except the one whose coefficients are all 0. For each such polynomial is of the form $f=\alpha+xf_1+yf_2$, where $\alpha\in F$ and f_1, f_2 are polynomials of lower degree than f. Suppose that f is the polynomial of least degree in two noncommuting indeterminates that vanishes for x and y. If $\alpha\neq 0$, then f_1, f_2 cannot both vanish; say $f_2\neq 0$, hence $\alpha x+xf_1x+yf_2x=0$, i.e., $x(f_1x+\alpha)=-yf_2x\neq 0$, a contradiction. Hence $\alpha=0$, so $xf_1=-yf_2$; by the choice of x and y this implies $f_1=f_2=0$, which contradicts the choice of f. So we have proved that there is no polynomial f other than the zero polynomial such that $f(x, y)=0$. In other words, the subalgebra generated by x and y is the *free associative algebra* on these generators. The restriction on the coefficients is easily lifted, so one has the following result ([23], [12], [25]):

THEOREM 4.3. *An entire ring is either a left and right Ore ring, or it contains a free algebra on two generators.*

By definition a *polynomial identity* is an identical relation not holding in all rings, in particular not in free rings. Thus Th.4.3 has the following immediate consequence [1]:

COROLLARY 1. *An entire ring with a polynomial identity is a (left and right) Ore ring.*

In analogy with Goldie's theorem, Posner [35] has generalized this result to show that any prime ring with a polynomial identity has a total ring of fractions which is a central simple algebra of finite dimension over its centre.

Jategaonkar who first proved Th.4.3 has also shown how to use it to embed the free algebra in a field [23]. Take a ring R which is a right but not left Ore ring. (E.g., $K[x;\ \alpha,\ 0]$ with a non-surjective endomorphism α, say $K=F(t)$ with $\alpha: f(t)\mapsto f(t^2)$.) By Th.4.3 R contains a free algebra and by Th.3.1, Cor.2 it has a field of (right) fractions. So the free algebra is embedded in a field. This is of interest because the free algebra is very far from being an Ore ring. However, this embedding is rather artificial; indeed most automorphisms of the free algebra cannot be extended to the field of fractions just constructed. Later, in Section 7, we shall meet fields of fractions which do not suffer from this defect.

The process of forming fractions can be applied to modules as well as to rings. If R is a ring with a subsemigroup S, and $\lambda: R\to R_S$ is the universal S-inverting homomorphism, then to each right R-module M there corresponds a right R_S-module M_S with an R-module homomorphism $\mu: M\to M_S$ (where M_S is regarded as R-module by means of $\lambda: x\cdot a = xa^\lambda$ for $x\in M_S$, $a\in R$), and μ is the universal mapping with this property, i.e., given any R-module homomorphism of M into an R_S-module N, there exists a unique R_S-module homomorphism of M_S into N such that the accompanying diagram commutes:

$$\begin{tikzcd} M \arrow[r] \arrow[dr] & M_S \arrow[d] \\ & N \end{tikzcd}$$

This much is general theory, proved in the same way as Th.2.1. There is even a formula for M_S if we are willing to use tensor products (cf. e.g., [27]):

$$M_S = M \otimes R_S,$$

but this formula makes it no easier to study M_S in detail. Now let us assume that S is a right denominator set in R; then the elements of M_S can be written as fractions m/s, where $m\in M$, $s\in S$, and two fractions represent the same element of M_S if and only if, over a suitable common denominator, they have the same numerator. The kernel of the natural mapping $\mu: M\to M_S$ is a submodule tM of M, called the *S-torsion submodule* of M. It consists of all $m\in M$ such that $ms=0$ for some $s\in S$. If $tM=0$, the module M is said to be *S-torsionfree*; e.g., the quotient M/tM is always S-torsionfree.

When R is the ring $\boldsymbol{Z}$ of integers and $S=\boldsymbol{Z}^*$ the set of all nonzero integers, tM reduces to the usual torsion subgroup of an abelian group.

5. Strongly regular rings. There have been many attempts to generalize the notion of 'inverse' of an element, to take account of zero-divisors, usually in the form of a 'relative inverse' a', satisfying $aa'a=a$. We shall present the part of this theory that is relevant to the embedding problem.

A ring R is said to be *regular* if to each $a\in R$ there corresponds $x\in R$ such that $axa=a$; if R is such that for each $a\in R$ there exists $x\in R$ satisfying $a^2x=a$, it is called *strongly regular*. In the commutative case this is the same as requiring R to be regular, but in general it is stronger. This is not apparent at first sight, but it will follow from the structure theorems given below. We shall need some of the standard theory of the Jacobson radical; this can be found in [22], to which we refer when necessary. Let us recall that from any family R_λ of rings we can form a *direct product* $P=\prod R_\lambda$ by taking the Cartesian (set-theoretical) product and performing all the operations componentwise. (In the older books this is also called the direct sum.) A subring R of the direct product P is said to be a *subdirect* product if the canonical projections ϵ_λ on the factors R_λ, when restricted to R, are still surjective. E.g., $\boldsymbol{Z}$ can be expressed as subdirect product of fields $\boldsymbol{Z}_p$, where p ranges over all primes. The projection of $\boldsymbol{Z}$ on $\boldsymbol{Z}_p$ maps each integer n to its residue class (mod p).

As an example of a strongly regular ring we have any field, or more generally, any direct product of fields. However, a subdirect product of fields need not be strongly regular, as the example of the integers shows. Nevertheless, these two notions are closely related:

THEOREM 5.1. *Every strongly regular ring is a subdirect product of fields.*

Proof. Let R be strongly regular. Then its Jacobson radical J is 0. For an element $a\in R$ lies in J precisely when $1-ax$ is invertible for all $x\in R$. By hypothesis we can find $x\in R$ such that $a(1-ax)=a-a^2x=0$ and $1-ax$ is invertible; hence $a=0$. It follows ([22], p. 14, [26], p. 58) that R is a subdirect product of primitive rings, each a homomorphic image of R and therefore again strongly regular, so it only remains to show that a strongly regular ring which is also primitive is a field. Now any primitive ring is a dense ring of linear transformations in a vector space V over a field ([22], p. 28, [26], p. 54), and we shall be done if we show that V is 1-dimensional. Assume that V contains two linearly independent elements v_1, v_2. By density there exists $a\in R$ such that $v_1a=v_2$, $v_2a=0$, and by strong regularity we can find $x\in R$ such that $a^2x=a$; hence $v_2=v_1a=v_1a^2x=v_2ax=0$, a contradiction. This completes the proof.

COROLLARY. *A ring is strongly regular if and only if it is regular and has no nilpotent elements other than* 0.

Proof. Let R be strongly regular; then it is a subdirect product of fields and therefore cannot have any nonzero nilpotent elements. Further, if $a^2x=a$, then

for each projection ϵ_λ of R on a factor K_λ of the product, either $a\epsilon_\lambda = 0$ or $a\epsilon_\lambda \cdot x\epsilon_\lambda = 1 = x\epsilon_\lambda \cdot a\epsilon_\lambda$. In all cases, $a\epsilon_\lambda \cdot x\epsilon_\lambda \cdot a\epsilon_\lambda = a\epsilon_\lambda$; hence $axa = a$ and R is regular.

Conversely, let R be regular without nonzero nilpotent elements, and take $x \in R$ to satisfy $axa = a$. Then

$$(a^2x - a)^2 = a^2xa^2x - a^2xa - a^3x + a^2 = a^3x - a^2 - a^3x + a^2 = 0,$$

hence $a^2x - a = 0$ and R is strongly regular.

The connection with fields of fractions is provided by the following result [7]:

THEOREM 5.2. *A subring of a strongly regular ring is embeddable in a field if and only if it is entire.*

Proof. The condition is clearly necessary; so assume that R is entire and is a subring of a strongly regular ring. The latter is a subdirect product of fields, so R is itself a subring of a direct product of fields, say

$$R \subseteq P = \prod K_\lambda.$$

Let $\epsilon_\lambda : P \to K_\lambda$ be the canonical projection and define for each $x \in P$,

$$\Gamma_x = \{\lambda \in \Lambda \mid x\epsilon_\lambda = 0\}, \qquad I_x = \prod_{\lambda \in \Gamma_x} K_\lambda.$$

Each I_x is an ideal in P. Let I be the ideal generated by all the I_x such that $x \in R^*$. Then $R \cap I = 0$; for if $x \in R \cap I$, then $x \in I_{y_1} + \cdots + I_{y_n}$, where $y_i \in R^*$, and hence $xy_1y_2 \cdots y_n = 0$. Therefore $x = 0$.

Let $f: P \to P/I$ be the natural homomorphism. Then by the construction of I, f is R^*-inverting. Since P/I, like P, is strongly regular, there is a homomorphism g of P/I into a field; now fg is an R^*-inverting homomorphism of P into a field, and (by Th.2.2) this provides an embedding of R in a field.

The following consequence was first proved using ultraproducts [36]:

COROLLARY. *An entire subring of a direct product of fields is embeddable in a field.*

Th.5.2 shows that an entire ring can be embedded in a strongly regular ring if and only if it can be embedded in a field. By contrast, any entire ring can be embedded in a regular ring; for there is a construction which associates with any ring R its 'total (right) quotient ring' $Q(R)$, and when R is entire (more generally, for any ring with 'zero singular ideal') $Q(R)$ is regular (cf. [17, 26]). Moreover, this total quotient ring agrees with the total ring of fractions when the latter exists, i.e., by Th.3.1, when the nonzero-divisors of R form a right denominator set. But in general the total quotient ring of R gives no clue about the embeddability of R in a field. E.g., though for an entire ring $Q(R)$ is always regular, it is not strongly regular unless R is a right Ore ring.

For a very thorough survey of quotient rings, see [40].

6. Topological embedding methods. Besides the time-honoured way of forming fractions there is another method of embedding rings in fields, which is also taught at school and goes back to the 16th century [38]. This is the method of decimal fractions; it consists of taking all expressions of the form

$$\sum_{-n}^{\infty} a_\nu t^\nu, \tag{1}$$

where $a_\nu = 0, 1, \cdots, 9$ and $t = 1/10$, and adding and multiplying in the usual way. Division is possible because if in (1), $a_{-n} \neq 0$ say, the series (1) can be written as $t^{-n}a_{-n}(1 - \sum_1^\infty b_\nu t^\nu)$, where each factor is invertible. Of course all the series are convergent, as Laurent series, because $|t| < 1$. This method can be generalized; the general form is even simpler in some respects, because the usual absolute value is replaced by a non-Archimedean valuation.

Let R be a ring with a *valuation*, i.e., a function $v(x)$ taking the integers or $+\infty$ as values, such that

V.1. $v(x) = \infty$ if and only if $x = 0$,

V.2. $v(xy) = v(x) + v(y)$,

V.3. $v(x - y) \geqq \min\{v(x), v(y)\}$.

Such a valuation may be thought of as defining a topology on R; since our aim is to embed R in a field (if possible), we shall specify the topology by its neighbourhoods of 1. We shall limit ourselves to the case where the set

$$P = \{ab^{-1} \mid a \in R, b \in R^*\} \tag{2}$$

is dense in the field to be constructed, so we must say when ab^{-1} is close to 1. The natural condition for this is to require $v(ab^{-1} - 1)$ to be large. Of course v is only defined on R in the first instance, but if we assume that it can be extended to the field of fractions in such a way as to satisfy V.1–3, we have

$$v(a - b) = v([ab^{-1} - 1]b) = v(ab^{-1} - 1) + v(b);$$

hence ab^{-1} is close to 1 precisely when $v(a - b) - v(b)$ is large.

We now have a topology (in fact a uniformity, cf. [6]) on the set P of fractions ab^{-1}. What is still needed to make it into a ring? We shall not make P itself into a ring, but its completion in the given topology. To enable us to add and multiply we need an "asymptotic Ore condition":

A. For any a, b in R^* the function

$$f(x, y) = v(ax - by) - v(by)$$

is unbounded above.

This enables us to embed R in a field, by defining the ring operations in the completion of P, much in the same way as the usual Ore condition was used before. The details are somewhat technical, so we omit them (cf. [8]), but there is a simple way of restating the result in terms of graded rings.

Let us define

$$R_n = \{x \in R \mid v(x) \geqq n\}.$$

Then the R_n form a descending series of additive subgroups of R such that $\cap R_n = 0$, $\cup R_n = R$ and $R_i R_j \subseteq R_{i+j}$. In other words, we have a *filtered ring*. With each such filtered ring R one associates another ring, its *graded ring* gr R, as follows: the additive group of gr R is the direct sum of the terms

$$\mathrm{gr}_n R = R_n / R_{n+1}. \tag{3}$$

To define multiplication it is enough, by the distributive law, to specify the product of an element of $\mathrm{gr}_i R$ and one of $\mathrm{gr}_j R$. Let $\alpha \in \mathrm{gr}_i R$, $\beta \in \mathrm{gr}_j R$; according to (3), these are cosets, say $\alpha = a + R_{i+1}$, $\beta = b + R_{j+1}$. We define $\alpha\beta = ab + R_{i+j+1}$; it is easily verified that this definition does not depend on the choice of a, b within their cosets. Associativity is clear, so gr R becomes a ring in this way. Loosely speaking, it is the ring formed by taking 'leading terms' in R.

We shall get a graded ring even if the function v satisfies, instead of V.2, only $v(xy) \geqq v(x) + v(y)$. The stronger condition V.2 merely ensures that gr R is entire; further the asymptotic Ore condition can be shown to be equivalent to the Ore condition for gr R. The result may be summed up as follows [8]:

THEOREM 6.1. *Let R be a filtered ring whose associated graded ring is a right Ore ring. Then R can be embedded in a field K. In fact K can be taken as a complete topological field, with P given by* (2) *as a dense subset.*

The result can be used to embed the universal associative envelope of a Lie algebra (even infinite-dimensional) in a field. In particular it provides another embedding of the free algebra, because this can be regarded as the universal associative envelope of the free Lie algebra. (Cf. [8] and for other applications [9].) For a further generalization see [42].

Another 'topological' method of constructing fields of fractions consists of taking an ordered group and forming 'Laurent series': Given a totally ordered group G and a commutative field F, consider the direct power F^G of F indexed by G. With each $f \in F^G$ we associate a subset $D(f)$ of G, its *support*, defined as

$$D(f) = \{s \in G \mid f(s) \neq 0\}.$$

E.g., the group algebra FG of G may be identified with the set of elements of finite support. In general it will not be possible to define the multiplication of F^G in such a way as to extend the operation on FG, for this requires that

$$fg = (\Sigma f(s)s)(\Sigma g(t)t) = \sum_u \left[\sum_{st=u} f(s)g(t) \right] u, \tag{4}$$

and here the inner sum on the right will generally contain infinitely many nonzero terms to be added. However, if both f and g have a *well-ordered* support, then the equation $st = u$ has, for a given $u \in G$, only finitely many solutions (s, t) in $D(f) \times D(g)$. Moreover, the sum (4) itself will then have well-ordered support. We can therefore define a ring structure on the set A of all elements of F^G whose support is well-ordered, in such a way that the group algebra FG becomes a subalgebra of A.

Finally it can be shown that A is in fact a field, so that the group algebra of G has been embedded in a field. This result was obtained simultaneously and independently by Malcev [**30**] and Neumann [**32**]. Later Higman [**21**] proved a general result on ordered algebraic systems which includes the Malcev-Neumann construction as a special case. For a simplified presentation of Higman's result see [**11**], p. 123.

We now have another way of embedding free algebras in fields: Let G be the free group on a set X. Then G can be totally ordered (by writing elements as products of basic commutators and taking the lexicographic ordering of the exponents, cf. [**31**]). Hence its group algebra FG is embeddable in a field. Since FG clearly contains the free algebra $F\langle X\rangle$ as subalgebra, this provides an embedding of the latter in a field.

7. The matrix method. In Section 2 we observed that to embed a noncommutative ring R in a field it may not be enough to produce inverses of all nonzero elements of R. We can try to overcome this difficulty by adjoining inverses of suitable matrices. Given a set Σ of square matrices over R, we can formally adjoin inverses of these matrices as follows. For each $n\times n$ matrix $A=(a_{ij})$ in Σ, take a set of n^2 symbols $A'=(a'_{ij})$ and adjoin the a'_{ij} to R as extra generators with defining relations, in matrix form,

$$AA' = A'A = I.$$

The resulting ring is denoted by R_Σ, and we have a natural homomorphism $\lambda: R\to R_\Sigma$. This ring has properties entirely analogous to the ring R_S described in Th.2.1, to which it reduces when all the matrices in Σ are 1×1. So we again call R_Σ and λ the *universal Σ-inverting ring* and *homomorphism*, respectively.

It is of interest to note that under suitable conditions on Σ, each element of R_Σ is some a'_{ij}; thus the generating set of R_Σ described above is then the whole ring. To state the result, let us say that the set Σ of matrices is *admissible* if (i) $1\in\Sigma$, (ii) the result of applying elementary row (or column) transformations to any matrix of Σ again lies in Σ, and (iii) if $A, B\in\Sigma$, then $\begin{pmatrix} A & C\\ 0 & B\end{pmatrix}\in\Sigma$ for any matrix C and zero matrix 0 of the appropriate size.

THEOREM 7.1. *Let R be a ring, Σ an admissible set of matrices over R, and $f: R\to S$ any Σ-inverting homomorphism. Then the set $\overline{R}$ consisting of all components of inverses of matrices in $\Sigma^{\&}$ is a subring of S.*

When S is a field, this result applies in particular to the set Σ_1 of all matrices over R whose images are invertible in S, for then Σ_1 can easily be shown to be admissible. When $\Sigma=\Sigma_1$, the set $\overline{R}$ is also called the *rational closure* of R under the homomorphism f.

The question now is: Which matrices do we have to invert to get a field? In the commutative case the answer was easy: we had to invert all nonzero elements, and this ensured that all matrices that are nonzero-divisors also become invertible. But in the general case there may well be matrices that are nonzero-divisors and yet are not invertible in any larger ring. Thus let R be any entire

ring that is neither a right nor a left Ore ring. Then there exist $a, b, c, d \in R^*$ such that $Ra \cap Rb = 0$, $cR \cap dR = 0$, and it follows easily that the matrix

$$\begin{pmatrix} ac & ad \\ bc & bd \end{pmatrix} = \begin{pmatrix} a \\ b \end{pmatrix} (c\ d) \tag{1}$$

is a nonzero-divisor. But this matrix cannot be invertible in any field; more generally, no homomorphism of R into a field can map (1) to an invertible matrix. This example suggests the following definition:

A matrix A over a ring R is said to be *full* if it is square, say $n \times n$, and it cannot be written as a product $A = PQ$, where P is $n \times r$, Q is $r \times n$, and $r < n$. Clearly, the most we can hope for, in mapping a ring R into a field, is to invert the full matrices. The next result goes some way towards saying when this can be done [**13**].

THEOREM 7.2. *Let R be a ring such that the set Φ of all full matrices over R is admissible. Then the universal Φ-inverting ring R_Φ is either* 0 *or a field; moreover, when R_Φ is a field, the universal Φ-inverting homomorphism $\lambda: R \to R_\Phi$ is injective.*

This is proved by exhibiting each element of R_Φ as a component of the solution of a matrix equation with a full matrix of coefficients, and showing that the inverse element satisfies a similar equation. The last part of the theorem follows from Th.2.2, because any nonzero element of a ring is full.

The hypotheses of Th.7.2 are satisfied if R_Φ is a nonzero ring in which each one-sided matrix inverse is two-sided (i.e., $AB = I$ implies $BA = I$, cf. [**14**]), but it is more difficult to find conditions in terms of R itself. Here is one case where this has been done.

A *free ideal ring*, or *fir* for short, is a ring R in which each right ideal (and each left ideal) is free as an R-module, and all bases of a free module have the same number of elements [**10**]. Examples of firs are: (i) free algebras over a commutative field (on any free generating set), (ii) group algebras of free groups, and (iii) free products of fields, over a common subfield, [**10**]. For a fir one can show that the set Φ of full matrices is admissible and that $\lambda: R \to R_\Phi$ is an embedding. Hence using Th.7.2 we see that each fir can be embedded in a field. Since the class of full matrices is preserved under automorphisms, each automorphism of the fir can be extended (in just one way) to an automorphism of its field of fractions.

A final point concerns the uniqueness. The field of fractions of an integral domain or, more generally, of a right Ore ring is unique up to isomorphism. For any ring isomorphism $a \mapsto a'$ extends to an isomorphism of the field of fractions by the formula $(a/s)' = a'/s'$. In the general case this is no longer so: there may be several nonisomorphic fields of fractions of a noncommutative ring. (E.g., for the free algebra, the field of fractions obtained from Th.7.2, using the fact that the free algebra is a fir, can be shown to be different from the field obtained by Jategaonkar's construction in Section 4.) Given two fields of fractions K_1, K_2 of a ring R, we define a *specialization* from K_1 to K_2 as a homomorphism f from a subring R_1 of K_1 to K_2 that reduces to the identity map on R and such

that any nonunit of R_1 is mapped to 0 by f. Of course no unit can be mapped to 0, so $\ker f$ is the precise set of nonunits. This means that the nonunits of R_1 form an ideal, $\mathfrak{m}$ say; a ring with this property is said to be *local*. Clearly $R_1/\mathfrak{m}$ is a field, isomorphic to the image of R_1 under f. This image is a subfield of K_2 containing R, but since K_2 is generated (as a field) by R, the image is all of K_2, i.e., f is surjective. This then shows each specialization to be surjective.

A field of fractions K of R is said to be *universal* if for each field of fractions K' of R there is a unique specialization from K to K'. This property determines K up to isomorphism, and in looking for fields of fractions, we are naturally interested in finding the universal one, if it exists. Any free algebra (over a commutative field) has a universal field of fractions [2] and also has other non-universal ones. More generally one can show [**14**]:

THEOREM 7.3. *Let R be a ring such that the set Φ of full matrices is admissible. If $R_\Phi \neq 0$ (so that R_Φ is a field, by Th.7.2), then R_Φ is the universal field of fractions of R.*

In particular this shows that each fir has a universal field of fractions.

These notions will be useful when one tries to do noncommutative algebraic geometry, which might be defined as the study of zero-sets of rational functions in skew fields, just as the usual kind is the study of zero-sets of polynomials in commutative fields. In the commutative case polynomial zero-sets and rational zero-sets are the same, which is why we could confine ourselves to the former. In general this may not be so (cf. the examples in [39]), and some thought is needed even to construct rational functions. In [2] Amitsur classified rational function fields according to the (rational) identities they satisfy; this is taken up by Bergman [39] from a more general view-point; he also shows that affine space over a field with infinite center is irreducible and describes a universal field of functions.

Clearly many problems remain; we end by listing a few:

1. Find criteria for the existence of a homomorphism of a ring into a field (remember that the zero-mapping is not a homomorphism).
2. Which rings have fields of fractions?
3. Which rings have more than one field of fractions?
4. Which rings have a universal field of fractions?

References

1. S. A. Amitsur, On rings with identities, J. London Math. Soc., 30 (1955) 464–470.

2. ———, Rational identities and applications to algebra and geometry, J. Algebra, 3 (1966) 304–359.

3. K. Asano, Über die Quotientenbildung von Schiefringen, J. Math. Soc. Japan, 1 (1949) 73–78.

4. L. A. Bokut', On Malcev's problem, Sibirsk. Mat. Zh., 10 (1969) 965–1005.

5. A. J. Bowtell, On a question of Malcev, J. Algebra, 6 (1967) 126–139.

6. N. Bourbaki, Topologie générale, Ch. I–II, Hermann, Paris 1949.

7. I. E. Burmistrovič, On the embedding of rings in fields, Sibirsk. Mat. Zh., 4 (1963) 1235–1240.

8. P. M. Cohn, On the embedding of rings in skew fields, Proc. London Math. Soc., (3) 11 (1961) 511–530.

9. ———, Quadratic extensions of skew fields, Proc. London Math. Soc., (3) 11 (1961) 531–556.

10. ———, Free ideal rings, J. Algebra, 1 (1964) 47–69.

11. ———, Universal Algebra, Harper & Row, (New York, London, Tokyo 1965).

12. ———, Free associative algebras, Bull. London Math. Soc., 1 (1969) 1–39.

13. ———, The embedding of firs in skew fields, to appear.

14. ———, Universal skew fields of fractions, to appear.

15. C. W. Curtis, A note on noncommutative polynomials, Proc. Amer. Math. Soc., 3 (1952) 965–969.

16. M.-L. Dubreil-Jacotin, Sur l'immersion d'un semigroupe dans un groupe, C. R. Acad. Sci. Paris, 225 (1947) 787–788.

17. C. C. Faith, Lectures on injective modules and quotient rings, Springer Lecture notes in Math. No. 49 (Berlin 1967).

18. A. W. Goldie, The structure of prime rings under ascending chain conditions, Proc. London Math. Soc., (3) 8 (1958) 589–608.

19. ———, Some aspects of ring theory, Bull. London Math. Soc., 1 (1969) 129–154.

20. R. Hartshorne, Residues and duality, Springer Lecture notes in Math., No. 20 (Berlin 1966).

21. G. Higman, Ordering by divisibility in abstract algebras, Proc. London Math. Soc., (3) 2 (1952) 326–336.

22. N. Jacobson, Structure of rings, Providence, R. I., 1956, 1964.

23. A. V. Jategaonkar, Ore domains and free algebras, Bull. London Math. Soc., 1 (1969) 45–46.

24. A. A. Klein, Rings nonembeddable in fields with multiplicative semigroups embeddable in groups, J. Algebra, 6 (1967) 101–125.

25. E. G. Koševoi, On certain associative algebras with transcendental relations, Algebra i Logika, Sem. 9, No. 5 (1970).

26. J. Lambek, Lectures on rings and modules, Blaisdell, Waltham, Mass., 1966.

27. S. MacLane, Homology, Springer, Berlin-New York, 1963.

28. A. I. Malcev, On the immersion of an algebraic ring into a field, Math. Ann., 113 (1937) 686–691.

29. ———, Über die Einbettung von assoziativen Systemen in Gruppen I, Mat. Sbornik, 6 (48) (1939) 331–336, II. dto 8 (50) (1950) 251–264.

30. ———, On the embedding of group algebras in division algebras, Dokl. Akad. Nauk SSSR, 60 (1948) 1499–1501.

31. B. H. Neumann, On ordered groups, Amer. J. Math., 71 (1949) 1–18.

32. ———, On ordered division rings, Trans. Amer. Math. Soc., 66 (1949) 202–252.

33. O. Ore, Linear equations in noncommutative fields, Ann. of Math., 32 (1931) 463–477.

34. ———, Theory of noncommutative polynomials, Ann. of Math., 34 (1933) 480–508.

35. E. C. Posner, Prime rings satisfying a polynomial identity, Proc. Amer. Math. Soc., 11 (1960) 180–184.

36. A. Robinson, A note on embedding problems, Fund. Math., 50 (1961/62) 167–309; 455–461.

37. E. Steinitz, Algebraische Theorie der Körper, J. Reine Angew. Math., 137 (1910) new edition by R. Baer and H. Hasse, Berlin, 1930.

38. S. Stevin, Arithmétique (Amsterdam 1585).

39. G. M. Bergman, Skew fields of noncommutative rational functions, after Amitsur, to appear.

40. V. P. Elizarov, Rings of fractions, Algebra i Logika Sem., 8, no. 4 (1969) 381–424.

41. A. A. Klein, A conjecture concerning embeddability of rings into fields, J. Algebra, (to appear).

42. J. Dauns, Embeddings in division rings, Trans. Amer. Math. Soc., 150 (1970) 287–299.

WHAT IS A CONVEX SET?

VICTOR KLEE, University of Washington

This is a slight expansion (and, of course, a translation) of the author's article [**63**] on *Convexité* in the new French encyclopedia, *Encyclopedia Universalis*, and appears here with the kind permission of the encyclopedia's publishers. Its purpose is to supply a broad but brief survey, at a rather elementary level, of several aspects of convexity theory. Those aspects are emphasized which appear to the author to be most active at present and to be most accessible to the chosen level of exposition. In the allotted space, the topics covered cannot, of course, be "surveyed" in the usual sense; instead, each is represented by one or more of its highlights. No proofs are included.

For theorems which are often designated in the literature by authors' names, the relevant names are given here, even though in some cases the implied attribution is incomplete. Beyond that, there has been no attempt at attribution. The references are in most cases not to the original or the most definitive sources, but rather to expository treatments or to papers which contain useful collections of additional references. While this policy probably maximizes the utility/size ratio of the bibliography, it has the unfortunate consequence of omitting the names of many prominent workers in the field. Those names can be found in the bibliographies of the references cited. A reference which appears in parentheses at the end of a paragraph contains useful information, or at least useful references, for the entire area of convexity theory with which the paragraph is concerned. In addition to items which are specifically mentioned in the text, the bibliography includes a number of books, monographs, lecture notes, symposium volumes, and survey articles in which the notion of convexity has played an important role. Most of these have appeared in the past fifteen years.

Preparation of this article was partially supported by the Office of Naval Research.

Introduction. The study of **convex sets** is a branch of geometry, analysis, and linear algebra that has numerous connections with other areas of mathematics and serves to unify many apparently diverse mathematical phenomena. It is also relevant to several areas of science and technology.

Victor Klee did his undergraduate work at Pomona College and went on to the Ph.D. at the University of Virginia under Prof. E. J. McShane. He was in the Faculty at Virginia for several years and spent a year at the Institute of Advanced Study before he joined the Faculty at the University of Washington. He has spent two years leave of absence at the University of Copenhagen and one year at UCLA; he received an Honorary Degree from Pomona College in 1965.He has done extensive research on convexity, combinatorics, functional analysis, topology and related fields. He served on the Council of the A.M.S. and has also devoted considerable energy to the M.A.A.; it suffices to note that he is the current President. In addition to numerous research papers, Professor Klee edited the A.M.S. Pure Mathematics Symposium Vol. 7, *Convexity* (1963) and was the translator of *Combinatorial Geometry in the Plane* by Hadwiger and Debrunner (HRW 1964). In my opinion he is one of the leading experts on convexity in the world. *Editor.*

Though convex sets are defined in various settings (see [27] for a survey), the most useful definitions are based on a notion of betweenness. When E is a space in which such a notion is defined, a subset C of E is called **convex** provided that for each two points x and y of C, C includes all points between x and y. The most important setting, and the only one to be discussed here, is that in which E is a vector space over the real number field R or, in particular, is the n-dimensional Euclidean space E^n, and the points between x and y are those of the line segment xy. Thus, a subset C of a real vector space is convex provided that C contains every segment whose endpoints both belong to C. (For example, a cube in E^3 is convex but its boundary is not, for the boundary does not contain the segment xy unless x and y lie together in some 2-dimensional face of the cube.) The importance of convexity theory stems from the fact that convex sets arise frequently in many areas of mathematics and are often amenable to rather elementary reasoning. Even the infinite-dimensional theory is based to a considerable extent on 2- and 3-dimensional reasoning.

The first systematic study of convexity was made by Minkowski (1864–1909), whose works [71] contain, at least in germinal form, most of the important ideas of the subject. The early developments of convexity theory were finite-dimensional and directed mainly toward the solution of quantitative problems; an excellent survey of them was made by Bonnesen and Fenchel [14] in 1934. Since 1940, however, the combinatorial, qualitative, and dimension-free parts of the theory have tended to predominate, perhaps because of their many applications in other areas of mathematics. After some preliminary material that is relevant to all parts of the theory, the present exposition begins with the quantitative and combinatorial aspects because they are restricted to the finite-dimensional spaces that are most likely to be familiar to the reader. In discussing, later, the qualitative and dimension-free aspects of the theory, some slight familiarity with topological vector spaces is assumed. The reader who lacks this familiarity may restrict his attention to the case of Euclidean n-space E^n.

A fascinating aspect of convexity theory is the large number of easily stated and intuitively appealing unsolved problems that it still contains. A few such problems are included here.

Preliminary Material. Any two distinct points x and y of a real vector space E determine a unique **line.** It consists of all points of the form $(1-\lambda)x+\lambda y$, λ ranging over all real numbers. Those points for which $\lambda \geqq 0$ and for which $0 \leqq \lambda \leqq 1$ form respectively the **ray** from x through y and the **segment** xy. An **affine** set is one that contains all lines determined by pairs of its points; equivalently, it is a translate of a linear subspace. For example, the affine sets in E^3 are the empty set, the onepointed sets, the lines, the planes, and E^3 itself. A **hyperplane** H in E is an affine set of deficiency or codimension 1; that is, H is not properly contained in any affine subset of E other than E itself. In particular, the hyperplanes of E^n are its affine subsets of dimension $n-1$. For any hyperplane H, the complement $E \sim H$ is expressible in a unique way as the union

of two convex sets. They are called the **open halfspaces** bounded by H and their unions with H are the **closed halfspaces** bounded by H. Two sets X and Y are said to be **separated** by H provided that X lies in one of these closed halfspaces and Y in the other. The set X is **supported** by H at the point x provided that x belongs to X but is separated from X by H. Fig. 1 shows a hyperplane H (in this case, a line) in E^2, separating the convex sets X and Y and supporting X at the point x.

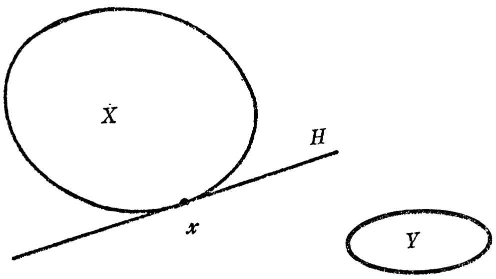


FIG. 1

Intimately related to the notion of a convex set is that of a convex function, which is important in most parts of convexity theory and in several areas of analysis. Let ϕ be a real-valued function whose domain D lies in a real vector space E. Then ϕ is called **convex** provided that D is convex and ϕ satisfies the inequality,

$$\phi((1-\lambda)x+\lambda y) \leqq (1-\lambda)\phi(x)+\lambda\phi(y),$$

for all points x and y of D and all numbers λ between 0 and 1. Equivalently, ϕ is a convex function if and only if its **epigraph** G is a convex set, where G is the subset of the product space $E \times R$ consisting of all ordered pairs (x, τ) such

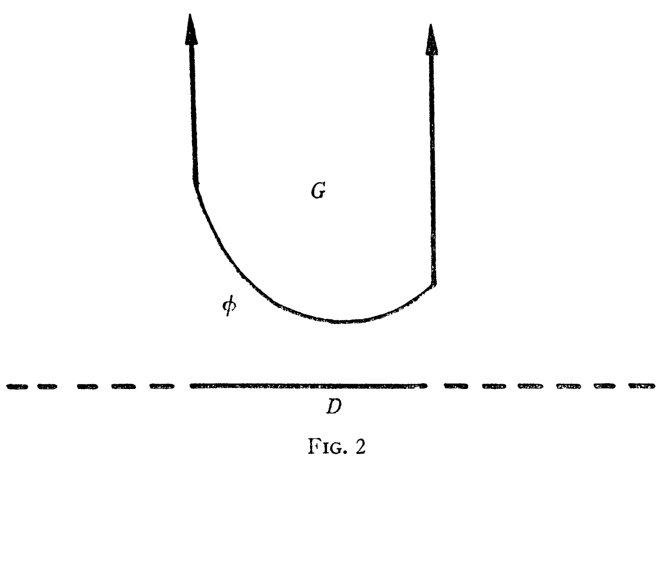


FIG. 2

that $x \in D$ and $\tau \geqq \phi(x)$. Fig. 2 shows a convex function ϕ, its domain D, and its epigraph G.

A function is called **concave** provided that its negative is convex, and **affine** provided that it is both convex and concave. A real-valued function on E is affine if and only if it differs by a constant from a linear function. The hyperplanes H in E are precisely the zero sets of the nonconstant affine functionals f on E. If H is the set of all x for which $f(x) = 0$, then the closed halfspaces bounded by H are determined by the inequalities $f(x) \leqq 0$ and $f(x) \geqq 0$.

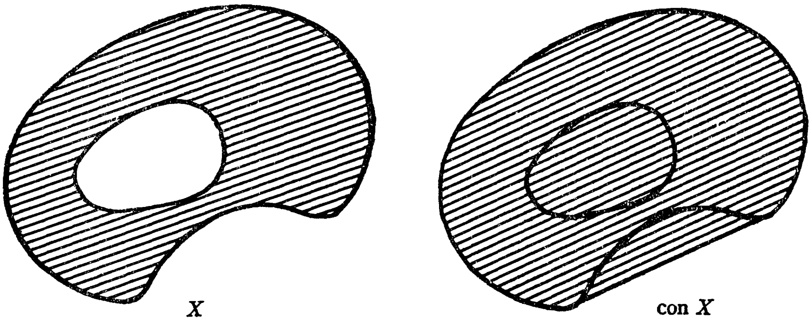


FIG. 3

The **convex hull** of a set X, denoted here by con X, is the intersection of all convex sets containing X. It is convex, as is any intersection of convex sets, and hence is the *smallest* convex set containing X. (Fig. 3 shows a nonconvex plane set and its convex hull.) Equivalently, con X is the set of all **convex combinations** of X—that is, points of the form $\sum_1^k \lambda_i x_i$, where the x_i's are points of X and the λ_i's are positive numbers whose sum is 1. For any such combination and for any convex function ϕ whose domain contains X,

$$\phi\left(\sum_1^k \lambda_i x_i\right) \leqq \sum_1^k \lambda_i \phi(x_i).$$

The **closed convex hull** of a set is the closure of its convex hull.

A point x of a convex set C is called an **extreme point** of C provided that $C \sim \{x\}$ is convex or, equivalently, that x is not an inner point of any segment in C. More generally, a **face** of C is a convex set $F \subset C$ such that F is not "crossed" by any segment in C—that is, $xy \subset F$ whenever x and y belong to C and F includes an inner point of xy. (For example, a cube in E^3 has six 2-faces, twelve 1-faces (edges), and eight 0-faces (extreme points). It is the convex hull of its set of extreme points.)

The terms *body* and *cone* are used in different ways by various authors. Here **body** always means a bounded closed subset of E^n that has nonempty interior, and **cone** means a set in a real vector space which is a union of rays from the origin O.

Quantitative Aspects. Convexity is a basic notion in the so-called geometry of numbers, and, indeed, it was the latter subject that led Minkowski to many of his investigations. One of his most striking results is that if C is a convex body in E^n which is symmetric about the origin O (that is, $x \in C$ implies $-x \in C$), and if the volume $V(C)$ is at least 2^n, then C includes at least one point other than O whose coordinates are all integers [21].

A **packing** of convex bodies is an arrangement in which no two of the bodies have common interior points. In addition to being of interest for themselves, packing problems are found in number theory, information theory, crystallography, botany, virology, and other areas of science. Often the interest is in packings of maximum density. The densest packing of congruent circular disks in E^2 is one in which the disks are inscribed in nonoverlapping regular hexagons covering E^2; each disk touches six others. (See Fig. 4.) It has long been conjectured that a densest packing of congruent spherical balls in E^3 is the cubic close-packing of Kepler, obtained by imagining the space to be divided into black and white cubes forming a 3-dimensional chessboard, and then placing a ball concentric with each black cube and tangent to each of the twelve edges of the cube; each ball touches twelve others. However, the conjecture has been proved only for packings of congruent balls whose centers form a lattice (if x and y are centers then so is $2x-y$). ([35] [81])

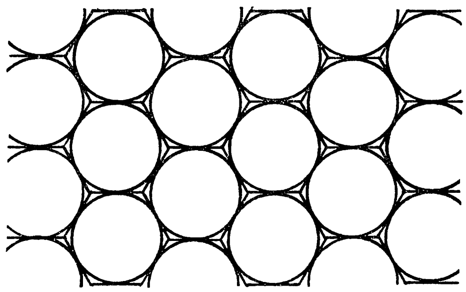


FIG. 4

There is a rich collection of quantitative results involving such measurements of convex bodies as volume, surface area, diameter, etc. Many extremal problems concerning such measurements have spherical balls or simplices as their solutions. (A **simplex** in E^n is the convex hull of $n+1$ points not contained in any hyperplane.) For example, the isoperimetric inequality asserts that if S and V are respectively the surface area and the volume of a convex body C in E^n, and if ω is the volume of an n-dimensional ball of unit radius, then $S^n \geqq n\omega V^{n-1}$, with equality if and only if C is a ball. Thus of all bodies with given volume, balls have the least surface area [13] [14] [46] [47]. Any convex body C of E^n lies in a unique ball of minimum radius r. Jung's inequality, of interest in approximation theory, asserts that if C's diameter is d then $r \leqq (n/(2n+2))^{1/2}d$, with equality if and only if C is a regular simplex [27] [47]. Loewner's theorem

asserts that any body in E^n lies in a unique ellipsoid of minimum volume [**28**].

Convex sets are prominent in the theory of geometric probability. For example, the following problem was posed (in a different form) by Sylvester. Let C be a convex body of unit volume in E^n and let $n+1$ points be chosen from C, independently and at random. Except in degenerate cases, the convex hull of these points is an n-simplex. What is the expected volume, V_C, of the simplex? For $n=2$, the values of V_C are between $1/12$ and $35/48\pi^2$, attained respectively when C is a triangle and when C is an ellipse. For larger values of n, V_C is known when C is a ball but not when C is a simplex [**61**].([**55**] [**72**][**73**])

For convex bodies $C_1, \cdots, C_k$ in E^n and positive numbers $\lambda_1, \cdots, \lambda_k$, the set of all points of the form $\lambda_1 x_1 + \cdots + \lambda_k x_k$ with $x_i \in C_i$ is another convex body C, denoted by $\lambda_1 C_1 + \cdots + \lambda_k C_k$. When $C_1, \cdots, C_k$ are fixed, the volume of C is expressible as a homogeneous polynomial of degree n in the parameters $\lambda_1, \cdots, \lambda_k$. Some of the deepest parts of the quantitative theory concern the coefficients of this polynomial, which are called the **mixed volumes** of $C_1, \cdots, C_k$. A basic tool in the study of mixed volumes is the Brunn-Minkowski theorem asserting that for $0<\lambda<1$, $(V((1-\lambda)C_1+\lambda C_2))^{1/n} \geqq (1-\lambda)(V(C_1))^{1/n}+\lambda(V(C_2))^{1/n}$ (that is, the nth root of the volume is a concave function of λ) and characterizing the cases of equality. Inequalities for mixed volumes yield the isoperimetric inequality and other inequalities of immediate geometric interest. ([**14**] [**46**] [**47**])

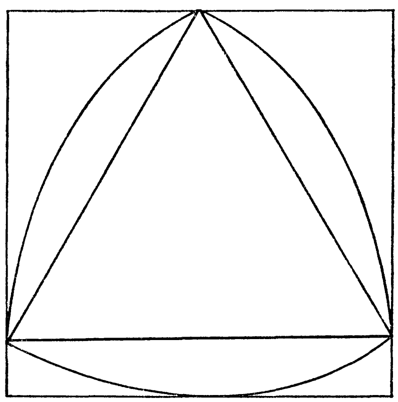


FIG. 5

A convex body C in E^n is said to be of **constant breadth** b provided that b is the distance between any two parallel supporting hyperplanes of C; equivalently, C is of diameter b and the diameter is increased by adding any point of E^n not in C. From the second description it follows that any set of diameter $\leqq b$ in E^n lies in at least one convex body of constant breadth b. Hence the following problem of Borsuk can be reduced to the case in which X is a convex body of constant breadth: Can every set X of diameter 1 in E^n be covered by

$n+1$ sets of diameter <1? The answer is affirmative for $n \leqq 3$, unknown for $n>3$ [42]. Noncircular plane convex bodies of constant breadth have been studied by many mathematicians. Their special properties have led to their use in kinematic linkages and in other mechanisms. Any such body can be placed in a square and then "rotated" while remaining in contact with all four sides of the square. (See Fig. 5.) ([5] [14] [87])

Having mentioned some unsolved problems in E^3 and E^4, we end this section with one in E^2. A **chord** of a convex body is a segment joining two boundary points, and an **equichordal point** is one through which all chords are of equal length. Does any plane convex body have two equichordal points? [60]

Combinatorial Aspects. Much of combinatorial convexity theory deals with intersection properties of convex sets. The intersection C of any family of convex sets is itself convex, though C may be empty. Helly's theorem asserts that C is nonempty if the convex sets are all in E^n, each $n+1$ of them have nonempty intersection, and the family is finite or its members are all compact. There are numerous generalizations and applications of Helly's theorem. From its 1-dimensional form it follows that if C is a cube in E^n then any family of pairwise intersecting translates of C has nonempty intersection. In fact, a convex body C has this intersection property if and only if C is affinely equivalent to a cube. ([27] [48] [87])

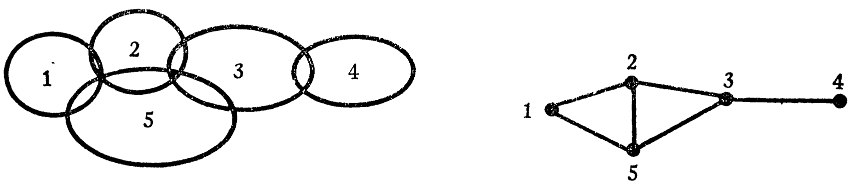


FIG. 6

The problem of determining all intersection properties of convex sets is not trivial even in E^1. It leads to the notion of an interval graph, which has been used in such diverse fields as molecular genetics, psychophysics, archaeology, and ecology. For any family of sets the associated **intersection graph** is an abstract graph having a node for each member of the family, two nodes being joined by an arc of the graph if and only if the corresponding sets intersect. Fig. 6 shows a family of convex sets and its intersection graph. Any finite graph can be realized as the intersection graph of a family of convex bodies in E^3, but not so in E^2 or E^1. An **interval graph** is one that is the intersection graph of a finite family of convex bodies in E^1. Such graphs have been characterized in various ways, but the corresponding problem relative to E^2 is still open. ([84] [86])

Another area of combinatorial research is concerned with the representation of convex hulls. The simplest and most useful result is Carathéodory's theorem, asserting that if $X \subset E^n$ and $u \in \text{con } X$ then $u \in \text{con } Y$ for some set Y consisting

of $n+1$ or fewer points of X. For example, when $X \subset E^2$ any point of con X belongs to X, to a segment determined by two points of X, or to a triangle determined by three points of X. Carathéodory's theorem has many generalizations and applications, including the fact that con X is compact for each compact $X \subset E^n$. ([27] [79])

The most extensive combinatorial developments deal with the facial structure of convex polyhedra. Though terminology has not been standardized, we here use the term **polyhedron** to mean a subset of E^n that is the intersection of a finite number of closed halfspaces. Of special interest are the bounded polyhedra, here called **polytopes,** and the polyhedral cones. By the lemma of Farkas, a set is a polytope if and only if it is the convex hull of a finite set of points, and is a polyhedral cone if and only if it is the convex hull of a finite number of rays from the origin. More generally, the following five conditions on a set P in E^n are equivalent: (i) P is a polyhedron; (ii) P is a closed convex set whose number of faces is finite; (iii) P is the convex hull of a finite system of points and rays; (iv) P is the vector sum $B+C=\{b+c\colon b\in B,\ c\in C\}$ of a polytope B and a polyhedral cone C; (v) P is the closed convex hull of the union of a polytope B and a translate of a polyhedral cone C. Fig. 7 shows a 2-dimensional polyhedron P and the associated sets B and C. ([56])

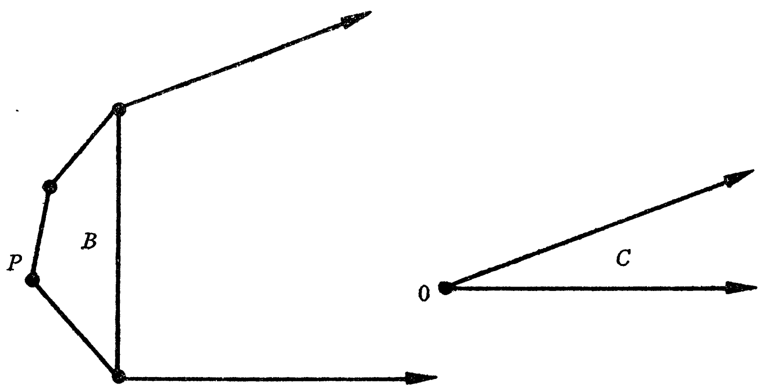


FIG. 7

In the combinatorial study of polyhedra, the first landmark was Euler's 1752 theorem asserting that $v-e+f=2$ for any 3-polytope, where v, e, and f are respectively the numbers of vertices, edges, and 2-faces. (Extreme points of polyhedra are usually called **vertices.**) The generalization of Schläfli and Poincaré asserts that if $f_i(P)$ is the number of i-dimensional faces of an n-dimensional polytope P, then $\sum_{i=0}^{n-1}(-1)^i f_i(P)=1-(-1)^n$. The second landmark was Steinitz's 1934 theorem characterizing the graphs of 3-polytopes, the combinatorial structures formed by vertices and edges, as those that are planar (representable in E^2 without crossings) and 3-connected (between any two vertices

there are three independent paths). The first graph of Fig. 8 corresponds to a cube. The second and third graphs of Fig. 8 do not correspond to any 3-polytope, for the first is not 3-connected and the second is not planar. However, the third graph does correspond to a 4-dimensional simplex. Various properties, including n-connectedness, have been established for the graphs of n-polytopes, but no combinatorial characterization is known for $n>3$. ([43])

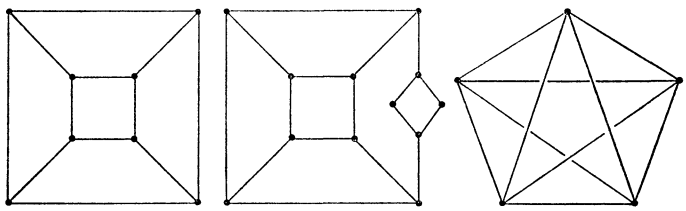


FIG. 8

A third landmark in the combinatorial study of polyhedra was the development, beginning in the late 1940's and still continuing, of computational techniques for minimizing linear functions on polyhedra. These techniques, known as *linear programming*, provide solutions to a wide range of practical optimization problems, and they are also useful for other computations involving polyhedra. Their importance led to a renewed interest in polyhedra and, for example, to rediscovery of the striking fact that for each $k>n$ there is an n-polytope with k vertices such that each $[n/2]$ vertices determine a face. A closely related development is the recent proof [70] that the maximum number of vertices possessed by any n-polyhedron with k $(n-1)$-faces is

$$\binom{k-[(n+1)/2]}{k-n}+\binom{k-[(n+2)/2]}{k-n}. \qquad \text{([26] [43] [45] [59])}$$

Qualitative Aspects. The topics to be discussed in this section include normed vector spaces, polarity, separation and support theorems, extreme point theorems, and fixed point theorems.

For a real-valued function ϕ on a real vector space E, any two of the following conditions imply the third: subadditivity ($\phi(x+y)\leqq\phi(x)+\phi(y)$ for all x, $y\in E$); positive homogeneity ($\phi(\lambda x)=\lambda\phi(x)$ for all $x\in E$ and $\lambda\geqq 0$); convexity. A **norm** is a function that satisfies these conditions as well as being symmetric ($\phi(-x)=\phi(x)$) and positive ($x\neq 0$ implies $\phi(x)>0$). A **normed vector space** consists of a vector space E together with a norm on E. The norm is usually denoted by $\|\ \|$ and leads to a useful notion of distance by defining the distance between x and y as $\|x-y\|$. The **unit ball** of such a space is the set of all points x for which $\|x\|\leqq 1$, while the **unit sphere** is defined by the condition $\|x\|=1$. When $r=2$, the function $\|x\|_r=(\sum_1^n |x_i|^r)^{1/r}$ is the usual Euclidean norm for E^n; an inequality of Minkowski asserts that it is a norm for all $r\geqq 1$. In general, the notion of convexity plays a key role in the theory of inequalities, and many useful

inequalities assert merely that a certain function is convex. Another important norm for E^n is given by

$$\|x\|_\infty = \max_{1\leq i\leq n} |x_i|.$$

The unit ball for $\| \ \|_\infty$ is an n-dimensional cube and for $\| \ \|_1$ is a so-called cross-polytope (a regular octahedron when $n=3$). Fig. 9 shows the unit spheres in E^2 associated with $\| \ \|_r$ for $r=1$, $3/2$, 2, 3, and ∞. ([29])

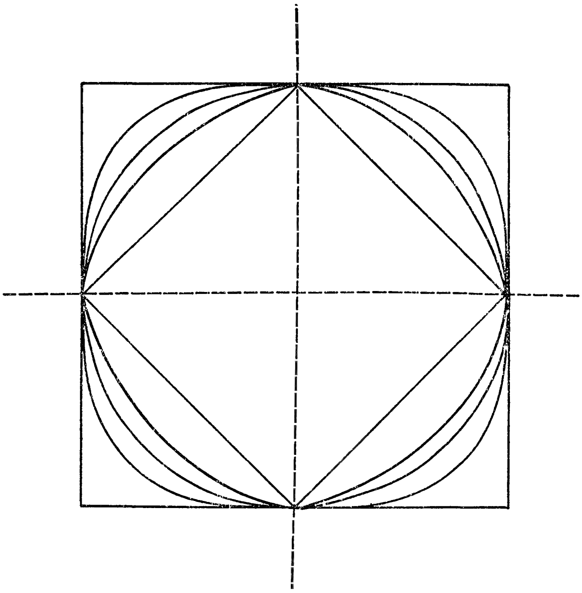


FIG. 9

The study of normed vector spaces is in a sense equivalent to the study of a certain class of convex sets. The equivalence involves the notion of the **gauge function** μ of a set U, where, for each $x\in E$, $\mu(x)$ is defined as the infimum of all numbers $\lambda>0$ such that $x\in\lambda U$. For any norm $\| \ \|$ on E, the associated unit ball U is a convex set that intersects every line through O in a closed segment having O as its midpoint; further, $\| \ \|$ is the gauge function of U. Conversely, for any convex set U of the sort described, the associated gauge function is a norm for which U is the unit ball. Thus any property of a normed vector space can be described completely in terms of its unit ball or its unit sphere. For example, the **rotundity** of a normed space may be defined by saying that $\|x+y\|<\|x\|+\|y\|$ whenever x and y are not collinear with O, or by saying, equivalently, that the unit sphere does not contain any line segments. **Smoothness** of the space E may be defined by saying that for any two points x and y not equal to 0, the function $\phi(\lambda)=\|x+\lambda y\|$ is differentiable at $\lambda=0$—or by saying, equivalently, that the unit sphere has at each point a unique supporting hyperplane. Euclidean n-space is both rotund and smooth. The behavior of a normed space can sometimes be

improved by **renorming,** which means introducing a new norm that is caught between two positive multiples of the original norm and hence induces the same topology on the space. For example, any separable normed linear space can be renormed so as to be simultaneously smooth and strictly convex. ([**24**] [**25**] [**29**])

All of the quantitative problems mentioned earlier for Euclidean spaces have been studied also for finite-dimensional normed spaces, commonly called **Minkowski spaces.** The analogue of Jung's inequality asserts that if E is an n-dimensional Minkowski space and X is a set of diameter d in E, then there is a point z of E such that $\|x-z\|\leqq(n/(n+1))d$ for all $x\in X$ [**27**]. In contrast to the Euclidean case, the unit ball of a Minkowski space need not be an extreme body so far as the isoperimetric problem for that space is concerned [**18**]. There are many characterizations of those Minkowski spaces which are equivalent to Euclidean spaces or, in other words, whose unit balls are ellipsoids. Most of these characterizations are related to the fact that, among the Minkowski spaces E of dimension $\geqq 3$, each of the following conditions characterizes the Euclidean spaces: (a) (Jordan—vonNeumann) $\|x+y\|^2+\|x-y\|^2=2\|x\|^2+2\|y\|^2$; (b) (Blaschke-Kakutani) for any hyperplane H through O in E, there is a linear projection of norm 1 of E onto H—equivalently, there is a line L such that if U is the unit ball then the intersection $U\cap H$ is equal to the intersection of U with the "cylinder" $U+L$. ([**29**] [**58**])

So far as the applications of convexity in other parts of mathematics are concerned, separation and support theorems are of special importance. They are widely used in functional analysis and have been used in game theory, in the theory of summability, and even to prove certain coloring theorems of graph theory. Together with Lyapunov's theorem asserting the convexity of the range of a nonatomic vector-valued measure [**9**] [**49**] [**68**], they are among the principal abstract tools of the theory of optimal control [**62**]. Separation theorems set forth conditions under which two nonempty disjoint convex subsets X and Y of a topological vector space E can be separated by a hyperplane, either in the weak sense defined above or in various stronger senses. It suffices, for example, that E should be finite-dimensional or that one of the sets should have nonempty interior. A consequence is that if C is a convex set whose interior is empty and A is a nonempty affine set disjoint from the interior of C, then A lies in a hyperplane separating A from C. In particular, a closed convex set C with nonempty interior is supported at each of its boundary points (Mazur-Bourgin). Though that conclusion may fail if C's interior is empty, the support points of C are dense in C's boundary if E is a Banach space (Bishop-Phelps) and also if E is locally convex and C weakly compact. If A is an affine subset of a real vector space E, ϕ is a convex function on E, and f is an affine function on A such that $f\leqq\phi$ on A, then f can be extended to an affine function g on E with $g\leqq\phi$ on E. This result, a slight improvement of the classical Hahn-Banach theorem of functional analysis, follows from the separation theorem applied to the graph of f and the epigraph of ϕ. Because of this and other relationships, separation and support

theorems may be regarded as geometric relatives of the Hahn-Banach theorem. ([16] [29] [54] [62] [64] [80] [83])

The notion of polarity is essential in convexity theory and in the theory of topological vector spaces. Let E and F be two spaces that are paired by a bilinear form $\langle\ ,\ \rangle$. For example, let $E=F=E^n$ and let $\langle\ ,\ \rangle$ be the usual inner product given by

$$\langle (x_1, \cdots, x_n), (y_1, \cdots, y_n)\rangle = x_1y_1 + \cdots + x_ny_n.$$

For any $X\subset E$ the **polar** X^0 of X is the set of all $y\in F$ such that $\langle x, y\rangle \leqq 1$ for all $x\in X$. The polar is always convex, being an intersection of halfspaces. In geometry this notion plays two roles that are dual to each other. When unable to prove directly a theorem about sets $X_1, X_2, \cdots$, one may find it possible to prove an equivalent statement about $X_1^0, X_2^0, \cdots$. On the other hand, having proved an interesting theorem about $X_1, X_2, \cdots$, one may find that the polar form of the theorem is also of interest. For $1\leqq r$ and $1/r+1/s=1$, the $\|\ \|_r$ unit ball of E^n is polar to the $\|\ \|_s$ unit ball of E^n. Some polar pairs of this sort appeared in Fig. 9. There are close relationships among the notion of the polar of a convex set, the so-called support function of a convex body, and the notion of the conjugate of a normed linear space. ([16] [29] [43] [54] [64])

Both convex and concave functions occur often in practical optimization problems, and both have properties that are helpful in such problems. Let f be a continuous real-valued function whose domain D lies in a locally convex topological vector space E. If f is convex, then any local minimum x_0 for f is a global minimum [80] [83] [88]. That is to say, if there is a neighborhood U of x_0 such that $f(x_0)\leqq f(x)$ for all $x\in U\cap D$, then the same inequality holds for all $x\in D$. This justifies various iterative procedures for finding or approximating x_0. [88] If f is concave and D compact, then f attains a minimum at an extreme point of D. That is one of the reasons for the importance of extreme points in functional analysis. The other reason is contained in theorems of Krein and Milman, asserting that if C is a compact convex subset of a locally convex space, and if $X\subset C$, then C is equal to the closed convex hull of X if and only if the closure of X includes all extreme points of C. Thus C's extreme points form the smallest set by means of which, using convex combinations, all points of C can be approximated. There are extensions of the Krein-Milman theorem to certain noncompact sets, and some sharpening is possible in the finite-dimensional case. For example, if C is a closed convex set in E^n and C contains no line, then C is the convex hull of its extreme points together with its extreme rays. (An **extreme ray** of C is a ray not "crossed" by any segment.) ([3] [6] [22] [29] [54] [64] [76])

When C is a compact convex set in a locally convex space E and X is the set of all extreme points of C, it follows from the Krein-Milman theorem that each point p of C is the barycenter of a probability measure μ carried by the closure $\bar{X}$ of X—that is,

$$f(p) = \int_{x\in\bar{X}} f(x)d\mu(x)$$

for all continuous linear functionals f on E. Several of the integral representation theorems of analysis are consequences of this. However, when X is not closed, it is desirable to have the sharper representation afforded by a measure that is carried by X rather than $\overline{X}$. Choquet's theorem, which has stimulated much research in recent years, asserts that such a representation is always possible when C is metrizable. Uniqueness of the representation, for all $p \in C$, is associated with a useful dimension-free notion of simplex. These ideas have been applied in several fields—for example, in potential theory and in the theory of operator algebras. ([3] [22] [41] [76])

In conclusion, we turn briefly to fixed-point theorems for convex sets, stating two of the simplest but most important ones. Both have been extended in many ways. The Brouwer-Schauder-Tychonov theorem asserts that if C is a compact convex set in a locally convex space and if ϕ is a continuous mapping of C into C, then there is at least one point p of C such that $\phi(p)=p$. The theorem and its relatives are used in many ways, such as proving existence theorems for differential and integral equations, minimax theorems for game theory, and various geometric properties of convex sets. (For example, a compact subset of E^n or of Hilbert space H is convex if and only if each point of the space admits a unique nearest point in the set. It is unknown whether "compact" may be replaced by "closed" in H, though it may be in E^n.) A recent computational development [66] makes it possible to regard this fixed-point theorem as a tool for *constructing* solutions of various sorts of systems, rather than merely establishing their existence. ([15] [30])

The Markov-Kakutani theorem asserts that if C is a compact convex subset of a topological vector space and if Φ is a commuting family of continuous affine transformations of C into C, then there is at least one point $p \in C$ such that $\phi(p)=p$ for all $\phi \in \Phi$. It is used to prove the existence of invariant means on commutative groups, of a finitely additive translation-invariant extension of Lebesgue measure to all bounded subsets of E^n, and in many other ways. Note that local convexity is not required for the Markov-Kakutani theorem. It is unknown whether the assumption can be abandoned in the case of the Krein-Milman extreme point theorem and the Tychonov fixed-point theorem. ([15] [16] [30])

References

1. A. D. Alexandrov, Die innere Geometrie der konvexen Flächen, Akademie-Verlag, Berlin, 1955. (Translated from the 1948 Russian edition.)

2. ———, Konvexe Polyeder, Akademie-Verlag, Berlin, 1958. (Translated from the 1950 Russian edition by H. Abetz, under the scientific editorship of W. Süss.)

3. E. Alfsen, Convex Compact Sets and Boundary Integrals, Springer, Berlin, 1971.

4. R. V. Benson, Euclidean Geometry and Convexity, McGraw-Hill, New York, 1966.

5. L. Beretta and A. Maxia, Insiemi convessi e orbiformi, Univ. Roma e Ist. Naz. Alta Mat. Rend. Mat. (5) 1 (1940) 1–64.

6. C. Berge, Espaces Topologiques: Fonctions Multivoques, Dunod, Paris, 1959.

7. C. Berge and A. Ghouila-Houri, Programming, Games, and Transportation Networks, Methuen, London and Wiley, New York, 1965. (Translated by M. Merrington and C. Ramanujacharyulu from the 1962 French edition.)

8. W. Blaschke, Kreis und Kugel, Teubner, Leipzig, 1916. (Reprinted in 1949 by Chelsea, New York.)

9. E. Bolker, A class of convex bodies, Trans. Amer. Math. Soc., 145 (1969) 323–345.

10. V. G. Boltyanskiĭ and I. Ts. Gohberg, Theorems and Problems of Combinatorial Geometry, (Russian) Nauka, Moscow, 1965.

11. V. G. Boltyanskiĭ and I. M. Yaglom, Convex figures and bodies (Russian), Encyclopedia of Elementary Mathematics, Nauka, Moscow, 1966, pp. 182–269.

12. V. G. Boltyanskiĭ and I. M. Yaglom, Geometric problems on maxima and minima (Russian), Encyclopedia of Elementary Mathematics, Nauka, Moscow, 1966, pp. 270–348.

13. T. Bonnesen, Les Problèmes des Isopérimétriques et des Isépiphanes, Gauthier-Villars, Paris, 1928.

14. ——— and W. Fenchel, Theorie der konvexen Körper, Springer, Berlin 1934. (Reprinted by Chelsea, New York, 1948.)

15. F. Bonsall, Lectures on Some Fixed Point Theorems of Functional Analysis, Tata Institute of Fundamental Research, Bombay, 1962.

16. N. Bourbaki, Espaces Vectoriels Topologiques (rev. ed.), Hermann, Paris, 1966, Chaps. 1–2 and 1967, Chaps. 3–5.

17. M. Brückner, Vielecke und Vielfläche, Teubner, Leipzig, 1900.

18. H. Busemann, The foundations of Minkowskian geometry, Comment. Math. Helv., 24 (1950) 156–187.

19. ———, Convex Surfaces, Interscience, Wiley, New York, 1958.

20. ———, The Geometry of Geodesics, Academic Press, New York, 1955.

21. J. W. S. Cassels, An Introduction to the Geometry of Numbers, Springer, Berlin, 1959.

22. G. Choquet, Lectures on Analysis (ed. by J. Marsden, T. Lance, and S. Gelbart), Benjamin, New York, 1969.

23. H. S. M. Coxeter, Regular Polytopes, Pitman, London, 1948. (Second ed. Macmillan, New York and Collier-Macmillan, London, 1963.)

24. D. F. Cudia, Rotundity, in Convexity, Proc. Sympos. Pure Math., Amer. Math. Soc., 7 (1963) 73–97.

25. ———, The geometry of Banach spaces. Smoothness, Trans. Amer. Math. Soc., 110 (1964) 284–314.

26. G. B. Dantzig, Linear Programming and Extensions, Princeton University Press, Princeton, 1963.

27. L. Danzer, B. Grünbaum, and V. Klee, Helly's theorem and its relatives, in Convexity, Proc. Sympos. Pure Math., Amer. Math. Soc., 7 (1963) 101–180. (Also, an expanded version of this, translated into Russian by S. Zalgaller under the editorship of I. M. Yaglom, published as Helly's Theorem and its Applications (Russian), *Mir*, Moscow, 1968.)

28. L. Danzer, D. Laugwitz, and H. Lenz, Über das Löwnersche Ellipsoid und sein Analogon unter den einem Eikörper einbeschriebenen Ellipsoid, Arch. Math. (Basel), 8 (1957) 214–219.

29. M. M. Day, Normed Linear Spaces, Springer, Berlin, 1958.

30. N. Dunford and J. T. Schwartz, Linear Operators, Part I: General Theory, Interscience, New York, 1958.

31. V. Eberhard, Zur Morphologie der Polyeder, Teubner, Leipzig, 1891.

32. H. G. Eggleston, Convexity, Cambridge University Press, New York, 1958.

33. ———, Problems in Euclidean Space: Application of Convexity, Pergamon, New York, 1957.

34. K. Fan, Convex Sets and their Applications, Lecture Notes, Argonne National Laboratory, Illinois, 1959.

35. L. Fejes Tóth, Lagerungen in der Ebene, auf der Kugel und im Raum, Springer, Berlin, 1953.

36. ———, Regular Figures, Pergamon, London, 1964.

37. W. Fenchel, ed., Convex Cones, Sets and Functions, Lecture Notes, Princeton University, 1951.

38. W. Fenchel, Proceedings of the Colloquium on Convexity, Copenhagen, Denmark, 1965, University of Copenhagen, 1967.

39. D. Gale, The Theory of Linear Economic Models, McGraw-Hill, New York, 1960.

40. A. A. Goldstein, Constructive Real Analysis, Harper and Row, New York, 1967.

41. A. Goullet de Rugy, Géométrie des Simplexes, Centre de Documentation Universitaire et S.E.D.E.S. Réunis, Paris, 1968.

42. B. Grünbaum, Borsuk's problem and related questions, in Convexity, Proc. Sympos. Pure Math., vol. 7, Amer. Math. Soc., (1963) 271–284.

43. ———, Convex Polytopes, Interscience-Wiley, London, 1967.

44. ——— and V. Klee, Convexity and Applications (L. Durst, ed.), Proc. CUPM Geometry Conference, Santa Barbara, 1967, MAA Committee on the Undergraduate Program in Mathematics, Berkeley, 1967.

45. ——— and G. C. Shepard, Convex polytopes, Bull. London Math. Soc., 1 (1969) 257–300.

46. H. Hadwiger, Altes und Neues über konvexe Körper, Birkhäuser, Basel, 1955.

47. ———, Vorlesungen über Inhalt, Oberfläche und Isoperimetrie, Springer, Berlin, 1957.

48. ———, H. Debrunner, and V. Klee, Combinatorial Geometry in the Plane, Holt, Rinehart, and Winston, New York, 1964.

49. H. Halkin, On the necessary condition for optimal control of nonlinear systems, J. Analyse Math, 12 (1964) 1–82.

50. A. D. Joffe and V. M. Tikhomirov, Duality of convex functions and extremum problems, Russian Math. Surveys, 23 no. 6 (1968) 53–124. (Translated by T. Garrett from Uspehi Mat. Nauk 23, no. 6 (1968) 51–116.)

51. S. Karlin, Mathematical Methods and Theory in Games, Programming, and Economics, vols. 1–2, McGraw-Hill, New York, 1960.

52. ——— and W. J. Studden, Tchebycheff Systems: with Applications in Analysis and Statistics, Interscience-Wiley, New York, 1966.

53. N. D. Kazarinoff, Geometric Inequalities, Random House, New York, 1961.

54. J. L. Kelley, I. Namioka, et al., Linear Topological Spaces, Van Nostrand, Princeton, N. J., 1963.

55. M. G. Kendall and P. A. P. Moran, Geometrical Probability, Hafner, New York, 1963.

56. V. Klee, Some characterizations of convex polyhedra, Acta Math., 102 (1959) 79–107.

57. ———, ed., Convexity, Proc. Symp. Pure Math., Amer. Math. Soc., 7 (1963).

58. ———, Review of a Paper by W. Rudin and K. T. Smith, Math. Rev., 23A (1962) 376-377.

59. ———, Convex polytopes and linear programming, Proceedings of the IBM Scientific Computing Symposium on Combinatorial Problems, IBM Data Processing Division, White Plains, N. Y., 1966, pp. 123–158.

60. ———, Can a plane convex body have two equichordal points? this MONTHLY, 76 (1969) 54–55.

61. ———, What is the expected volume of a simplex whose vertices are chosen at random from a given convex body? this MONTHLY, 76 (1969) 286–288.

62. ———, Separation and support properties of convex sets—a survey, Control Theory and the Calculus of Variations, Academic Press, New York, 1969, pp. 235–303.

63. ———, Convexité, Encyclopaedia Universalis, Paris, 4 (1970) 982–985.

64. G. Köthe, Topological Vector Spaces I, Springer, New York, 1969. (Translated from the 1965 German edition by D. J. H. Garling.)

65. M. A. Krasnoselskiĭ and Ya B. Rutickiĭ, Convex Functions and Orlicz Spaces, Noordhoff, Groningen, 1961. (Translated by L. F. Boron from the 1958 Russian edition.)

66. H. Kuhn, Approximate Search for Fixed Points, Computing Methods in Optimization Problems—2, Academic Press, New York, 1969, pp. 199–211.

67. ——— and A. W. Tucker, eds., Linear Inequalities and Related Systems, Princeton University Press, 1956.

68. J. Lindenstrauss, A short proof of Liapounoff's convexity theorem, J. Math. Mech., 15 (1966) 971–972.

69. L. A. Lyusternik, Convex Figures and Polyhedra, Dover, New York, 1963. (Translated by T. J. Smith from the 1956 Russian edition.)

70. P. McMullen, The maximum numbers of faces of a convex polytope, Mathematika, 17 (1970) 179–184.

71. H. Minkowski, Gesammelte Abhandlungen, Teubner, Berlin, 1911.

72. P. A. P. Moran, A note on recent research in geometrical probability, J. Appl. Probability, 3 (1966) 453–463.

73. ———, A second note on recent research in geometrical probability, Advances in Appl. Probability, 1 (1969) 73–89.

74. J. J. Moreau, Fonctionnelles Convexes, Lecture Notes, Séminaire "Equations aux Dérivées Partielles," College de France, 1966–67.

75. T. S. Motzkin, Beiträge zur Theorie der linearen Ungleichungen, Ph.D. Thesis, Basel, 1933 (Jerusalem, 1936).

76. R. R. Phelps, Lectures on Choquet's Theorem, Van Nostrand, Princeton, N. J., 1966.

77. V. Pogorelov, Die Verbiegung konvexer Flächen, Akademie-Verlag, Berlin, 1958. (Translated from the 1951 Russian edition.)

78. T. Popoviciu, Les Fonctions Convexes, Hermann, Paris, 1945.

79. J. R. Reay, Generalizations of a theorem of Carathéodory, Memoir No. 54, Amer. Math. Soc., 1965.

80. R. T. Rockafellar, Convex Analysis, Princeton University Press, 1970.

81. C. A. Rogers, Packing and Covering, Cambridge University Press, 1964.

82. E. Steinitz and H. Rademacher, Vorlesungen über die Theorie der Polyeder, Springer, Berlin, 1934.

83. J. Stoer and C. Witzgall, Convexity and Optimization in Finite Dimensions I, Springer, Berlin, 1970.

84. A. C. Tucker, Matrix characterizations of circular-arc graphs, Pacific J. Math., 35 (1971) to appear.

85. F. A. Valentine, Convex Sets, McGraw-Hill, New York, 1964.

86. G. Wegner, Eigenschaften der Nerven homologisch-einfacher Familien im R^n, Ph.D. Thesis, Göttingen, 1967.

87. I. M. Yaglom and V. G. Boltyanskiĭ, Convex Figures, Holt, Rinehart, and Winston, New York, 1961.

88. W. Zangwill, Nonlinear Programming: A Unified Approach, Prentice-Hall, Englewood Cliffs, N. J., 1969.

CYCLIC PURSUIT OR "THE THREE BUGS PROBLEM"

M. S. KLAMKIN, Ford Motor Company, and D. J. NEWMAN, Yeshiva University

I. Introduction. A well-known problem that keeps making the rounds is the one of three (or more) bugs pursuing each other cyclically, each traveling with the same speed and having started initially at the vertices of a regular polygon. Usually one wants to know the distance traveled by each bug until mutual capture. The problem of the three bugs can be traced back to H. Brocard [1] in 1877. However, it appears probable that a wide dissemination of the problem was first due to the problem (with four and three dogs instead of bugs) appearing in the book of Steinhaus [2] which first appeared in 1950.

The case of four bugs is the simplest. By symmetry, the four bugs always remain on four vertices of a square (albeit of decreasing side). Since bug B_2 is always moving at right angles to the pursuing bug B_1, the speed of closure for these two bugs is just the speed of B_1. Consequently, the distance traveled from start until mutual capture is unity for unit speed and unit initial square. The case for n bugs is only slightly more difficult. Again by symmetry, the n bugs are always on n vertices of a regular n-gon. Here we resolve the velocity of B_2 along

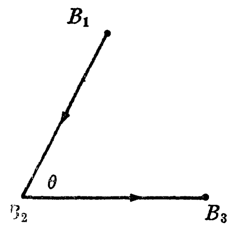


FIG. 1

and perpendicular to the velocity of B_1. Then the speed of closure is (see Fig. 1)

$$v = 1 + \cos\theta = 1 - \cos 2\pi/n,$$

and the distance traveled as well as the time to capture are both $1/v$ or $2^{-1}\csc^2 \pi/n$ which monotonically increases with n.

Another associated problem (which also has appeared a number of times) is to determine the paths of each of the bugs. Approximations to the paths can be obtained by replacing the relevant system of non-linear differential equations by their difference equation analogues. This can be done completely geometrically as follows in Fig. 2. A small increment Δ is chosen, and we lay off points

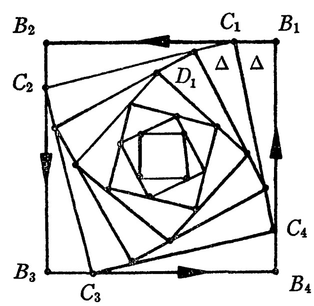


FIG. 2

C_1, C_2, C_3, C_4 on the edges of the initial square such that $B_1C_1 = B_2C_2 = B_3C_3 = B_4C_4 = \Delta$. We then repeat the same procedure starting from the square $C_1C_2C_3C_4$ and so on. The path of bug B_1 will then be approximately given by the envelope of the segments B_1C_1, C_1D_1, $\cdots$. The accuracy of the approximation will depend on the size of Δ. If we consider the paths for n bugs and let

$n \to \infty$, we generate a vortex. Figure 3 illustrates the paths for $n=6$. These figures are reminiscent of the mathematical themes of designs by R. Boyd which appeared from time to time in early issues of *Scripta Mathematica*

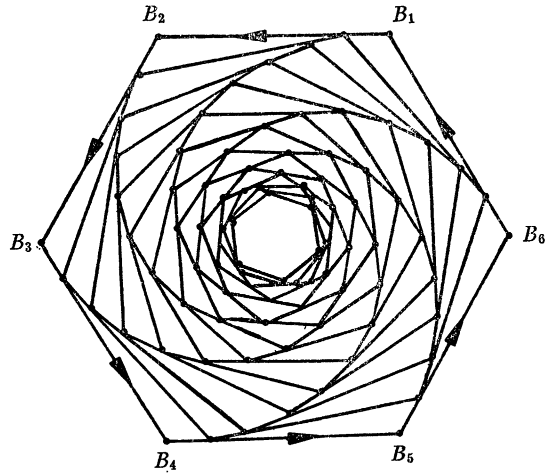


FIG. 3

One of the analytical solutions for the bug paths [3] was obtained by exploiting the symmetry of the initial configuration and then using polar coordinates. We shall give another simple solution via complex numbers. This will lead to a generalization of the three bug problem in which the bugs have different speeds, which are functions of the initial configuration which, in turn, need not be regular. This will lead to some interesting geometrical consequences which, incidentally, are associated with the Brocard points of the triangle.

A more difficult three bugs problem was first posed to one of us by Leo Moser about fifteen years ago. He conjectured that the three bugs, each moving with the same speed, could meet only simultaneously even if the initial configuration was an arbitrary triangle (but non-degenerate). This same problem was raised in [4], where additionally there is a solution of the n bug symmetrical case. After proving this conjecture, one of us found a reference to the problem and a partial solution while browsing through an out-of-print elementary text on ordinary differential equations by H. Bateman [5]. Bateman attributes the problem to Professor Morley and notes that it is treated numerically by F. E. Hackett (*Johns Hopkins Circular*, July, 1908). Our solution, while strategically the same as Bateman's, appears to be simpler in tactics. Additionally, Bateman does not rule out the case of mutual capture in an infinite length of time.

II. The Symmetrical n-Bugs Problem. Here we have n bugs $B_1, B_2, \cdots, B_n$ starting initially from the n vertices of a regular n-gon and pursuing each other cyclically with unit speeds. Let $B_r(t)$, $r=1, 2, \cdots$, be complex numbers denoting the position of the bug B_r, where the origin is taken to be the centroid of the

initial polygonal configuration. The equations of motion are then

$$\dot{B}_i = \frac{B_{i+1} - B_i}{|B_{i+1} - B_i|}, \qquad i = 1, 2, \cdots, n, \tag{1}$$

where $B_{n+1}=B_1$, $\sum_{i=1}^n B_i=0$, and $\dot{B}=dB/dt$.

All the previous solutions for this symmetrical case that we have seen are incomplete in that although existence is established by actually finding a solution, uniqueness is not demonstrated. Most likely, the uniqueness was tacitly assumed by virtue of the physical model. Uniqueness of the system (1) will follow [6] by showing that the real and imaginary parts of the r.h.s. satisfy a Lipschitz type condition of order 1 (for $B_r \neq B_s$), i.e.,

$$\left| \frac{p}{\sqrt{p^2+q^2}} - \frac{r}{\sqrt{r^2+s^2}} \right| \leq k\{ |p-r| + |q-s| \}$$

(which can be shown).

The usual assumption that the n bugs will always form a regular polygon is equivalent to

$$B_{i+1}(t) = \omega B_i(t) \qquad (\omega = e^{2\pi i/n}). \tag{2}$$

Since this assumption leads to a valid solution, which is unique, the assumption is justified. Equation (2) enables us to uncouple the system (1), i.e.,

$$\dot{B}_i = \frac{B_i(\omega - 1)}{|B_i|\,|\omega - 1|}.$$

Now letting $B_i = re^{i\theta}$, we obtain $\dot{r} = -\sin \pi/n$, $r\dot{\theta} = \cos \pi/n$. Whence,

$$r = r_0 - t \sin \pi/n,$$

$$\theta = \theta_0 - \left\{\cot \frac{\pi}{n}\right\} \log \frac{r_0 - t \sin \pi/n}{r_0},$$

$$r = r_0 \exp\{(\theta_0 - \theta) \tan \pi/n\}.$$

III. The Unsymmetrical 3-Bugs Problem with Different Speeds. In this generalization of the preceding case, the bugs start at the vertices of an arbitrary triangle (nondegenerate) with appropriate speeds such that they always form a similar triangle (Fig. 4). Since the closing speed for AB is $v_a + v_b \cos B$, etc., the speeds must satisfy

$$\frac{v_a + v_b \cos B}{c} = \frac{v_b + v_c \cos C}{a} = \frac{v_c + v_a \cos A}{b} = \lambda'.$$

Whence,

$$v_a = \frac{\lambda'(c - a \cos B + b \cos B \cos C)}{1 + \cos A \cos B \cos C} = \lambda(c - a \cos B + b \cos B \cos C).$$

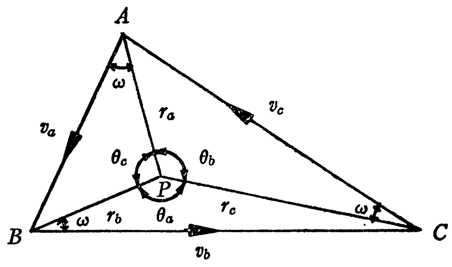


FIG. 4.

Of course v_b and v_c are obtained cyclically from v_a. In order to have cyclic pursuit, the three speeds must be positive. By the law of cosines,

$$v_a = \frac{4\lambda\Delta^2}{a^2c} = kb/a, \qquad v_b = kc/b, \qquad v_c = ka/c,$$

where Δ denotes the area of ABC.

The equations of motion are now

$$\dot{A} = \frac{v_a(B-A)}{|B-A|}, \qquad \dot{B} = \frac{v_b(C-B)}{|C-B|}, \qquad \dot{C} = \frac{v_c(A-C)}{|A-C|}, \tag{3}$$

where A, B, C are complex numbers measured from a point P which will be appropriately determined. We assume a point P exists such that from P the configuration of the three bugs appears to remain the same for all time. This requires

$$B = \frac{r_b}{r_a} Ae^{i\theta_c}, \qquad C = \frac{r_c}{r_a} Ae^{-i\theta_b}. \tag{4}$$

Then,

$$\begin{aligned} \dot{A} &= v_a \frac{A}{|A|} \frac{(r_be^{i\theta_c} - r_a)}{|r_be^{i\theta_c} - r_a|}, \\ \dot{B} &= v_b \frac{A}{|A|} \frac{(r_ce^{-i\theta_b} - r_be^{i\theta_c})}{|r_ce^{-i\theta_b} - r_be^{i\theta_c}|}. \end{aligned} \tag{5}$$

Since $\theta_a + \theta_b + \theta_b = 2\pi$,

$$\frac{r_av_b(r_ce^{i\theta_a} - r_b)}{|r_ce^{i\theta_a} - r_b|} = \frac{r_bv_a(r_be^{i\theta_c} - r_a)}{|r_be^{i\theta_c} - r_a|}, \tag{6}$$

and cyclically. Comparing magnitudes, we must have

$$r_a/v_a = r_b/v_b = r_c/v_c. \tag{7}$$

Also, if ΔPAB is not to change shape for small displacements of A and B, then to first order terms

$$\measuredangle PAB = \measuredangle PBC = \measuredangle PCA = \omega. \tag{8}$$

Both conditions (7) and (8) are satisfied if P is one of the two Brocard points [7] of the triangle. If the direction of the cyclic pursuit were reversed, with an appropriate change in the speeds, then P would be the other Brocard point. These two points are related to each other, being isogonal conjugates.

We can now complete the verification of (4). Since P is a Brocard point,

$$\frac{r_a}{b} = \frac{\sin\omega}{\sin A}, \qquad \frac{r_b}{c} = \frac{\sin\omega}{\sin B}, \qquad \frac{r_c}{a} = \frac{\sin\omega}{\sin C},$$

where ω is the Brocard angle. Also, $\theta_a = \pi - C$, $\theta_b = \pi - A$, $\theta_c = \pi - B$. Thus, (5) reduces to

$$\arg\{a^2 b e^{iC} + c^2 a\} = \arg\{c^2 a e^{iB} + b^2 c\},$$

or

$$\frac{a^2 b \sin C}{a^2 b \cos C + c^2 a} = \frac{c^2 a \sin B}{c^2 a \cos B + b^2 c},$$

which follows from the laws of sines and of cosines.

We now determine A from (4). But first

$$\begin{aligned} \frac{r_b e^{i\theta_c} - r_a}{|r_b e^{i\theta_c} - r_a|} &= -\frac{(ce^{iB}\sin A + b\sin B)}{|ce^{iB}\sin A + b\sin B|} \\ &= -\frac{(a^2+b^2+c^2-4i\Delta)}{2(a^2b^2+b^2c^2+c^2a^2)^{1/2}} = e^{i(\alpha+\pi/2)}. \end{aligned}$$

Letting $A = re^{i\theta}$, we get $\dot{r} + ir\dot{\theta} = ike^{i\alpha}b/a$. Whence,

$$\begin{aligned} r &= r_0 - \frac{kbt}{a}\sin\alpha, \\ \theta &= \theta_0 - \{\cot\alpha\}\log\frac{r_0 - (kbt\sin\alpha)/a}{r_0}, \\ r &= r_0\exp\{(\theta_0 - \theta)\tan\alpha\}. \end{aligned}$$

Now B and C can be determined from (4). We let $a = b = c = 1$; then this solution reduces to the solution for the symmetrical case determined previously.

It is to be noted that the paths would remain the same if we changed the three speeds v_a, v_b, v_c to $v_aF(t)$, $v_bF(t)$, $v_cF(t)$, where $F(t) > 0$. This only changes the time scale, i.e.,

$$t \to \int_0^t F(x)\,dx.$$

This dependence of the path only on the ratios of the speed was noted in [3].

IV. The Unsymmetrical 3-Bugs Problem with the Same Speeds. Here the 3 bugs each travel at the same speed, but the initial configuration is an arbitrary triangle (nondegenerate). We shall show that the meeting of the three bugs must be mutual, and it occurs in a finite time. Figure 4 will also apply here, but with $v_a = v_b = v_c = 1$. Three of the differential equations of the motion are

$$\dot{a} + 1 + \cos C = \dot{b} + 1 + \cos A = \dot{c} + 1 + \cos B = 0. \tag{9}$$

To find the rate of change of angles, we differentiate $2bc \cos A = b^2 + c^2 - a^2$ to obtain $-bc\dot{A} \sin A = b\dot{b} + c\dot{c} - a\dot{a} - (b\dot{c} + \dot{b}c) \cos A$. We now eliminate $\dot{a}$, $\dot{b}$, $\dot{c}$ using (9). Then by using the identity $b(\sin A \sin B - \cos A \cos B + 1) = c(\cos A - \cos B) + a(1 + \cos C)$, we obtain

$$A = \frac{\sin A}{b} - \frac{\sin B}{c}, \qquad B = \frac{\sin B}{c} - \frac{\sin C}{a}, \qquad C = \frac{\sin C}{a} - \frac{\sin A}{b}. \tag{10}$$

Alternatively, we could let the bugs take allowable small displacements and calculate the angles of the new triangle, using first order terms only.

Our proof is now indirect. Assume that $c \to 0$ as $t \to t_0^-$, while a and b do not $\to 0$. Since $\dot{a}$, $\dot{b}$, $\dot{c} \leqq 0$, we have a, b, c decreasing monotonically and thus a, b also approach a limit. So as $t \to t_0^-$, we have

$$c \to 0, \qquad a \to a_0 > 0, \qquad b \to b_0 > 0. \tag{11}$$

For t sufficiently close to t_0, we have $a/b - b/c < 0$. Then by the law of sines, $(\sin A)/b - (\sin B)/c < 0$. It now follows from (10) that A is eventually decreasing. Also, for all $t < t_0$, the angles lie in $(0, \pi)$ and this is essential in the proof. For if any angle ever became π, then by the uniqueness for our system of differential equations, it would have had to be always equal to π, which is contrary to the initial conditions. Whence,

$$\lim_{t \to t_0^-} A \quad \text{exists and is } < \pi. \tag{12}$$

We next prove that

$$\lim_{t \to t_0^-} B = 0. \tag{13}$$

From $a > a_0$ and (10), we have

$$B - \frac{\sin B}{c} > -\frac{1}{a_0}. \tag{14}$$

Thus, $\dot{B} > -1/a_0$, so that $B + t/a_0$ is increasing. Since it is bounded by $\pi + t_0/a_0$, the limit $\lim B = B_0$ exists for $t \to t_0^-$. Integrating (14), we obtain

$$B_0 - B(0) - \int_0^{t_0} \frac{\sin B}{c} \, dt > -\frac{t_0}{a_0},$$

so that the latter integral converges. If we now regard c as the independent variable (we may since it is monotone), we may use (9) to change the integral to

$$\int_0^{c(0)} \frac{\sin B}{1+\cos B} \frac{dc}{c} = \int_0^{c(0)} \tan\frac{B}{2}\frac{dc}{c}.$$

We already know that B_0 exists, so that the integrability of $(\tan B/2)/c$ around 0 insures that $\tan B_0/2=0$. This proves (13).

Finally, we recall from (11) that eventually c is the smallest side. Hence C is eventually the smallest angle, and so from (13) we get

$$\lim_{t\to t_0^-} C = 0. \tag{15}$$

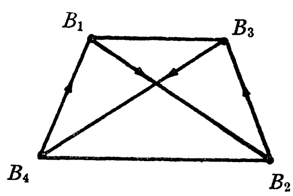
Relations (12), (13), and (15) give us a contradiction since $A+B+C=\pi$. Consequently, if there is a capture, it must be mutual. We now show that the time to capture is finite. From (9), we have

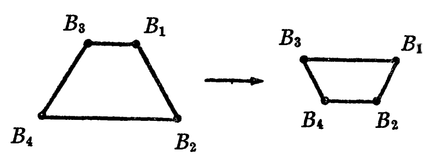
$$\dot a + \dot b + \dot c = -(3 + \cos A + \cos B + \cos C).$$

Since $1<\cos A+\cos B+\cos C\leqq 3/2$, we have $-9/2\leqq \dot a+\dot b+\dot c\leqq -4$ and $-9t/2 \leqq a+b+c-a(0)-b(0)-c(0)\leqq -4t$. Thus the time of capture t_c satisfies

$$\frac{a(0)+b(0)+c(0)}{9/2} \leqq t_c < \frac{a(0)+b(0)+c(0)}{4}.$$

V. An Open n-Bug Problem. It appears as if the previous mutual capture for 3-bugs may also be valid for n-bugs in any dimension, provided that initially all the bugs are not collinear. At first glance it may appear as if the following example of 4-bugs whose initial configuration is an isosceles trapezoid is a counter-example:





Since the trapezoid starts to flatten out, one may be tempted to guess that the bugs will become collinear and thus lead to a nonmutual capture. But collinearity can occur in only two ways. The first way is that B_1 captures B_2 the same time B_3 captures B_4 (by symmetry). The second is that the four bugs are

all separated. The latter case cannot happen since, by uniqueness, if the bugs ever became collinear they always were collinear, which violates the initial conditions. The former case is still open. Note that here we cannot use the previous uniqueness argument since the Lipschitz conditions do not hold at any capture.

Note added in proof. We are grateful to A. W. Walker for calling our attention to the following set of interesting and extensively referenced papers of A. Bernhart on pursuit problems, all appearing in SCRIPTA MATHEMATICA:

1. Curves of pursuit, 20 (1954) 125–141.
2. Curves of pursuit-II, 23 (1957) 49–65.
3. Polygons of pursuit, 24 (1959) 23–50.
4. Curves of general pursuit, 24 (1959) 189–206.

In (3), Bernhart notes that our problem III was set in the Camb. Math. Tripos Exam. of 1871 by R. K. Miller. Bernhart gives a solution different than ours and also considers the case of n bugs.

We are also grateful to N. L. Laurance for programming the nonsymmetric 4-bugs problem on a P.D.P. computer with a scope for visual output of the motion of the 4-bugs. The visual output indicated that the motion is one of mutual capture as indicated in the figures in V.

References

1. Nouv. Corres. Math., 3 (1877) pp. 175, 280.
2. H. Steinhaus, Mathematical Snapshots, Oxford University Press, New York, 1969, p. 136.
3. A. Watton and D. W. Kydon, Analytical Aspects of the N-Bug Problem, Amer. Jour. Physics, 37 (1969) 220–221.
4. J. C. Clapham, Playful Mice, Recreat. Math. Mag., Aug. 1962, pp. 6–7.
5. H. Bateman, Differential Equations, Longmans, Green and Co., London, 1918, pp. 8–10.
6. E. L. Ince, Ordinary Differential Equations, Dover, New York, 1944, p. 71.
7. R. A. Johnson, Advanced Euclidean Geometry, Dover, New York, 1960, pp. 263–270.

DEDEKIND-RADEMACHER SUMS

EMIL GROSSWALD, Temple University

Dedicated to the memory of Professor H. Rademacher

1. Introduction. In his study of two fragments on elliptic modular functions found among Riemann's posthumous papers, Dedekind [2] introduced the function

$$\eta(\tau) = \exp\left[\frac{\pi i \tau}{12}\right] \prod_{m=1}^{\infty} (1 - e^{2\pi i m \tau}),$$

known to-day as Dedekind's η-function. He proved (among many other results) that if a, b, c, d are integers satisfying $ad-bc=1$ and $c\neq 0$, and if

$$\tau' = \frac{a\tau + b}{c\tau + d},$$

then

$$\log \eta(\tau') = \log \eta(\tau) + \frac{1}{2}\log\frac{c\tau+d}{i} + \pi i\frac{a+d}{12c} - \pi i S. \tag{1}$$

In itself such a formula is, of course, trivial and may be considered as nothing more than a definition of $S=S(\tau, a, b, c, d)$. However, using a somewhat different notation, Dedekind proved the following far from trivial statements:

(i) *S is independent of τ;*

(ii) *S also does not depend on a or b;*

(iii) *$6cS$ is an integer.*

From this it follows in particular that $S=s(d, c)=A/B$, *with coprime integers* A, B, *and* $B\,|\,6c$.

Dedekind actually obtained an explicit representation for S. Let $[x]$ stand for the greatest integer not in excess of x, and denote by $((x))$ the function that vanishes on the integers and is equal to $x-[x]-\frac{1}{2}$ otherwise; then

$$s(d, c) = \sum_{\mu=0}^{c-1}\left(\left(\frac{d\mu}{c}\right)\right)\left(\left(\frac{\mu}{c}\right)\right). \tag{2}$$

These sums have since been called **Dedekind-Rademacher Sums.** Originally d and c were coprime, because they had to satisfy $ad-bc=1$; in the definition (2) no trace is left of any transformation $\tau\to(a\tau+b)/(c\tau+d)$, but it is still convenient (although not indispensable; see [2] and [9]) to keep this condition $(d, c)=1$.

By applying successively (1) to $\eta(\tau')$ and to $\eta(-1/\tau)$, and by observing that

$$\tau'' = \frac{a(-1/\tau)+b}{c(-1/\tau)+d} = \frac{b\tau - a}{d\tau - c},$$

Dedekind also proved the fundamental results:

Reciprocity Theorem. *For* $(c, d)=1$, *let* $s(d, c)$ *and* $s(c, d)$ *be defined by* (2); *then*

$$s(d, c) + s(c, d) = -\frac{1}{4} + \frac{1}{12}\left(\frac{c}{d} + \frac{1}{cd} + \frac{d}{c}\right) \tag{3}$$

holds.

Once (1) is known, the proof of (3) is fairly simple; however, the proof—or even the meaning—of (1) is rather sophisticated. Dedekind's own proof of (1) is not particularly difficult [2], and since then several other still neater proofs have been obtained [4, **11**]. Yet $s(d, c)$ is an elementary finite sum, and (3) is a simple arithmetic relation between two such sums; therefore, it is esthetically unsatisfactory to prove (3) in a round-about way by using properties of the transcendental function $\eta(\tau)$.

A direct, simple proof of (3), starting from the definition of $s(d, c)$, appears to be eminently desirable. The first such proof, due to H. Rademacher [6], ap-

peared in 1928, and since then many essentially distinct proofs have been found, several of them by H. Rademacher himself [7, 8, 10], others by L. Carlitz [1], L. J. Mordell [5], and others, e.g., [3] and [12].

It is the purpose of this paper to present what seems to be the simplest (although not the most elementary) proof of (3). There can be no claim of originality for the present proof. In fact, Rademacher gave a similar one [8] written in Hungarian (German summary); furthermore, in one of his posthumous manuscripts there is a hint that makes it extremely likely that he was in possession of the present treatment. Actually, the proof here presented is nothing more than an attempt to reconstruct from that hint what must have been Rademacher's own proof.

2. A Lemma. In order to make this presentation self contained, we include a proof of the following, well-known lemma [8]:

LEMMA. *For coprime, positive, rational integers c and d,*

$$s(d, c) = \frac{1}{4c} \sum_{m=1}^{c-1} \cot \frac{\pi m}{c} \cot \frac{\pi m d}{c} \cdot \tag{4}$$

As a first step, we show that, with $\zeta = e^{2\pi i/c}$, we have

$$\left(\left(\frac{\mu}{c}\right)\right) = \frac{1}{c} \sum_{n=1}^{c-1} \left(\frac{\zeta}{1-\zeta^n} + \frac{1}{2}\right) \zeta^{\mu n}. \tag{5}$$

It is clear from its definition that $((\mu/c))$ is a periodic function of period c; hence it has a finite Fourier series of the form

$$\left(\left(\frac{\mu}{c}\right)\right) = \sum_{m=0}^{c-1} a_m \zeta^{m\mu}. \tag{6}$$

In order to determine the Fourier coefficients a_m, we proceed routinely: We multiply both members of (6) by $\zeta^{-\mu n}$ and sum over μ:

$$\sum_{\mu=0}^{c-1} \left(\left(\frac{\mu}{c}\right)\right) \zeta^{-\mu n} = \sum_{\mu=0}^{c-1} \sum_{m=0}^{c-1} a_m \zeta^{\mu(m-n)} = \sum_{m=0}^{c-1} a_m \sum_{\mu=0}^{c-1} \zeta^{\mu(m-n)}.$$

The inner sum (a finite geometric progreisson) vanishes unless $m = n$, when its value is c, so that

$$a_n = \frac{1}{c} \sum_{\mu=0}^{c-1} \left(\left(\frac{\mu}{c}\right)\right) \zeta^{-\mu n}.$$

For $n = 0$, in particular, $a_0 = (1/c) \sum_{\mu=0}^{c-1} ((\mu/c)) = 0$, because the sum is equal to

$$\sum_{\mu=1}^{c-1} \left(\frac{\mu}{c} - \frac{1}{2}\right) = \frac{1}{c} \frac{(c-1)c}{2} - \frac{c-1}{2} = 0.$$

For $n\neq 0$,

$$a_n = \frac{1}{c}\sum_{\mu=1}^{c-1}\left(\frac{\mu}{c}-\frac{1}{2}\right)\zeta^{-\mu n} = \frac{1}{c^2}\sum_{\mu=1}^{c-1}\mu\zeta^{-\mu n} - \frac{1}{2c}\left(\sum_{\mu=0}^{c-1}\zeta^{-\mu n}-1\right)$$
$$= \frac{1}{c^2}\sum_{\mu=1}^{c-1}\mu\zeta^{-\mu n} + \frac{1}{2c},$$

because $\sum_{\mu=0}^{c-1}\zeta^{-\mu n}=0$. To compute the last sum, set $\zeta^{-n}=x$; then

$$\sum_{\mu=1}^{c-1}\mu x^{\mu} = x\sum_{\mu=1}^{c-1}\mu x^{\mu-1} = x\frac{d}{dx}\left(\sum_{\mu=0}^{c-1}x^{\mu}\right) = x\frac{d}{dx}\left(\frac{1-x^c}{1-x}\right)$$
$$= \frac{cx(1-x^{-1})}{(1-x)^2} = \frac{c}{x-1} = \frac{c}{\zeta^{-n}-1} = \frac{c\zeta^n}{1-\zeta^n},$$

where use has been made of $x^c=\zeta^{-nc}=1$. Consequently, for $n\neq 0$,

$$a_n = \frac{1}{c}\left(\frac{\zeta^n}{1-\zeta^n}+\frac{1}{2}\right),$$

and (5) is proved.

The proof of the Lemma is now immediate: By (2) and (5),

$$s(d,c) = \sum_{\mu=0}^{c-1}\left(\left(\frac{\mu}{c}\right)\right)\left(\left(\frac{\mu d}{c}\right)\right)$$
$$= \frac{1}{c^2}\sum_{\mu=0}^{c-1}\sum_{n=1}^{c-1}\left(\frac{\zeta^n}{1-\zeta^n}+\frac{1}{2}\right)\sum_{m=1}^{c-1}\left(\frac{\zeta^m}{1-\zeta^m}+\frac{1}{2}\right)\zeta^{(dm+n)\mu}$$
$$= \frac{1}{c^2}\sum_{n=1}^{c-1}\left(\frac{\zeta^n}{1-\zeta^n}+\frac{1}{2}\right)\sum_{m=1}^{c-1}\left(\frac{\zeta^m}{1-\zeta^m}+\frac{1}{2}\right)\sum_{\mu=0}^{c-1}\zeta^{(dm+n)\mu}.$$

The inner sum vanishes except for $dm+n\equiv 0 \pmod c$, when it equals c, so that

$$s(d,c) = \frac{1}{c}\sum_{m=1}^{c-1}\left(\frac{\zeta^m}{1-\zeta^m}+\frac{1}{2}\right)\left(\frac{\zeta^{-dm}}{1-\zeta^{-dm}}+\frac{1}{2}\right) = \frac{1}{4c}\sum_{m=1}^{c-1}\frac{1+\zeta^m}{1-\zeta^m}\cdot\frac{1+\zeta^{-dm}}{1-\zeta^{-dm}}.$$

Replacing here ζ by its value $e^{2\pi i/c}$, one obtains (4).

3. The Proof of the Reciprocity Theorem. Let $F(z)=\cot\pi z\cot\pi cz\cot\pi dz$, and consider the rectangle of vertices $\pm iM$, $1\pm iM$. The function $F(z)$ has poles at $z=0$ and $z=1$ on this contour; therefore if we want to integrate $F(z)$, we have to modify the contour by indentations at these points. We take as indentations identical "small" semicircles leaving $z=0$ inside, $z=1$ outside, and call the new contour $\mathfrak{C}$ (see Fig. 1).

Clearly, $F(z)=F(z+1)$, so that the integrals along the vertical sides (including the indentations) cancel each other. Also, $\lim_{M\to+\infty}\cot(x+iM)=-i$ uniformly for $0\leq x\leq 1$, so that $\lim_{M\to+\infty}F(x+iM)=(-i)^3=i$, and (similarly)

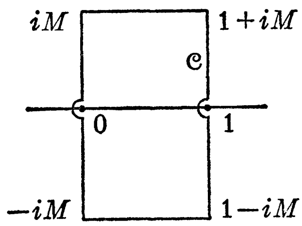


FIG. 1

$\lim_{M\to-\infty} F(x+iM) = -i$, both uniformly for $0 \leq x \leq 1$. Observe that $F(x+iy)$ is holomorphic for $0 < M \leq y \leq M_1$; it follows that $\int_{\mathfrak{C}} F(z)dz$ is in fact independent of M; consequently, $\int_{\mathfrak{C}} F(z)dz = \lim_{M\to\infty} \int_{\mathfrak{C}} F(z)dz = -2i$, and

$$\frac{1}{2\pi i}\int_{\mathfrak{C}} F(z) = -\frac{1}{\pi} = S, \tag{7}$$

where S stands for the sum of the residues of $F(z)$ at its singularities inside $\mathfrak{C}$. These singularities are (i) $z=0$, a triple pole, (ii) $z=\lambda/c$ ($\lambda = 1, 2, \cdots, c-1$), simple poles, (iii) $z=\mu/d$ ($\mu=1, 2, \cdots, d-1$), simple poles.

The poles at $z\neq 0$ are indeed all simple, because $\lambda/c=\mu/d$ is ruled out by $(c, d)=1$. The residues at the simple poles are

$$\frac{1}{\pi c}\cot\frac{\pi\lambda}{c}\cot\frac{\pi d\lambda}{c} \quad \text{and} \quad \frac{1}{\pi d}\cot\frac{\pi\mu}{d}\cot\frac{\pi c\mu}{d},$$

respectively. To find the residue at the triple pole, one uses the expansion $\cot z = (1/z) - (z/3) - \cdots$ and finds, in a neighborhood of $z=0$,

$$F(z) = \frac{1}{\pi^3 cdz^3}\left(1 - \frac{\pi^2z^2}{3} - \cdots\right)\left(1 - \frac{\pi^2z^2d^2}{3} - \cdots\right)\left(1 - \frac{\pi^2z^2c^2}{3} - \cdots\right),$$

so that the residue of $F(z)$ at $z=0$ is $-(1/3\pi)(c/d+1/cd+d/c)$. It follows that the sum of the residues of $F(z)$ inside $\mathfrak{C}$ is

$$S = -\frac{1}{3\pi}\left(\frac{c}{d}+\frac{1}{cd}+\frac{d}{c}\right) + \frac{1}{\pi c}\sum_{\lambda=1}^{c-1}\cot\frac{\pi\lambda}{c}\cot\frac{\pi d\lambda}{c} + \frac{1}{\pi d}\sum_{\mu=1}^{c-1}\cot\frac{\pi\mu}{d}\cot\frac{\pi c\mu}{d}\cdot$$

Using the Lemma to replace the last two sums, one obtains

$$S = \frac{4}{\pi}\left[-\frac{1}{12}\left(\frac{c}{d}+\frac{1}{cd}+\frac{d}{c}\right) + (s(c, d) + s(d, c))\right].$$

Substituting this in (7), after trivial simplifications, one has

$$s(c, d) + s(d, c) = -\frac{1}{4} + \frac{1}{12}\left(\frac{c}{d}+\frac{1}{cd}+\frac{d}{c}\right),$$

i.e., (3) and the proof of the Reciprocity Theorem is complete.

This paper was written with partial support from the National Science Foundation through the grant NSF GP-23170.

References

1. L. Carlitz, The reciprocity theorem for Dedekind sums, Pacific J. Math., 3 (1953) 523–527.

2. R. Dedekind, Erläuterungen zu den Fragmenten XXVIII, B. Riemann's Gesammelte Mathematische Werke, 2nd ed., (1892) 466–478; also Dedekind's Gesam. Math. Werke 1, (1930) 159–173.

3. V. Dieter, Zur Theorie der Dedekindschen Summen, Inaugural Dissertation, Kiel (1957) (mimeographed).

4. S. Iseki, The transformation formula for Dedekind's modular function and related functional equations, Duke Math. J., 24 (1957) 653–662.

5. L. J. Mordell, On the reciprocity formula for Dedekind sums, Amer. J. Math., 73 (1951) 593–598.

6. H. Rademacher, Zur Theorie der Modulfunktionen, Atti. del Congreso Intern. dei Mat. Bologna, (1928) 297–301.

7. ———, Zur Theorie der Modulfunktionen, J. Reine Angew. Math., 167 (1931) 312–366.

8. ———, Egy Reciprocitásképletröl a Modulfüggvények Elméletéböl, Mat. es Fiz. Lapok, 40 (1938) 24–43.

9. ———, and A. L. Whiteman, Theorems on Dedekind sums, Amer. J. Math., 63 (1941) 377–407.

10. ———, Die Reziprozitätsformel für Dedekindsche Summen, Acta Sci. Math. (Szeged), 12 (B) (1950) 57–60.

11. ———, On the transformation of log $\eta(\tau)$, J. Indian Math. Soc., 19 (1955) 25–30.

12. L. Rédei, Elementarer Beweis und Verallgemeinerung einer Reziprozitätsformel für Dedekindsche Summen, Acta Sci. Math. (Szeged), 12 (B) (1950) 236–239.

CORRECTION TO "THE JEEP ONCE MORE OR JEEPER BY THE DOZEN"

DAVID GALE, University of California at Berkeley

In the above article (this MONTHLY, 77 (1970) 493–501) I made the following statement (page 497, Line 6): "We note the familiar fact that a round trip can be substantially cheaper than two one-way trips." I am grateful to Professor R. B. Bruckel for pointing out to me (in a tactful manner) that this assertion is nonsense and a complete nonsequitur mathematically. The true state of affairs is just the other way around. On f loads of fuel, a single jeep can reach the point p (Equation (3)) where

$$p = d(f) = 1 + \frac{1}{3} + \cdots + \frac{1}{2f-1}\cdot$$

On $2f$ loads, the jeep can make a round trip of length $\bar{d}(2f)$ (Equation (6)), where

$$\bar{d}(2f) = 1 + \frac{1}{2} + \frac{1}{3} + \cdots + \frac{1}{2f}$$

and thus reach the point

$$p' = \frac{1}{2}\bar{d}(2f) = \frac{1}{2} + \frac{1}{4} + \frac{1}{6} + \cdots + \frac{1}{4f},$$

so

$$p - p' = \frac{1}{2}\left(1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \cdots + \frac{1}{2f-1} - \frac{1}{2f}\right).$$

Hence point p' falls short of point p by an amount approaching $\frac{1}{2}\log 2$, hence $\bar{d}(2f)$ falls short of $2d(f)$ by the distance $\log 2$. We may then ask how much additional fuel Δf is required to cover this additional distance. For this, we must solve

$$\log 2 = \frac{1}{2f+1} + \frac{1}{2f+2} + \frac{1}{2f+\Delta f}.$$

The sum on the right for large f is close to

$$\int_{2f}^{2f+\Delta f} \frac{1}{x}\,dx = \log((2f+\Delta f)/2f) = \log(1 + \Delta f/2f),$$

hence Δf is approximately $2f$, and we see that for large distances round trips are nearly four times as expensive as one-way trips!

Knowing the answer, it is now clear why this should be so! For a round trip, it is necessary to set up very substantial fuel depots at the far end of the desert, whereas for two one-way trips, the closer one gets to the far end of the desert, the smaller the depots have to be.

MATHEMATICAL NOTES

EDITED BY DAVID DRASIN

Manuscripts for this Department should be sent to Robert Gilmer, Department of Mathematics, Florida State University, Tallahassee, FL 32306.

FRACTIONAL DERIVATIVES AND LEIBNIZ RULE

T. J. OSLER, Rensselaer Polytechnic Institute

1. Introduction. The fractional derivative is an extension of the familiar derivative, $d^n f(z)/dz^n$, to nonintegral values of n. Fractional differentiation is of use in the solution of ordinary [6], partial [12], and integral equations [2, 3], as well as in other contexts, a few of which are indicated in the bibliography. Although other methods of solution are available, the fractional derivative approach to these problems often suggests methods that are not obvious in a clas-

$$\frac{\Gamma(c)\Gamma(c-a-b)}{\Gamma(c-a)\Gamma(c-b)} = F(a, b; c; 1).$$

In applying the Leibniz rule, u, v, and uv are restricted by the condition that the fractional derivatives occurring are defined by (1). This means that $\mathrm{Re}(c-a)$, $\mathrm{Re}(1-b)$, and $\mathrm{Re}(c-a-b)$ must be positive. It is well known [14, p. 281] that only the last of these three is necessary.

References

1. M. A. Bassam, Some properties of Holmgren-Riesz transform, Ann. Scoula Norm. Sup. Pisa (3), 15 (1961) 1–24.
2. A. Erdelyi, An integral equation involving Legendre functions, SIAM J. Appl. Math., 12 (1964) 15–30.
3. ———, Axially symmetric potentials and fractional integration, *ibid.*, 13 (1965) 216–228.
4. A. K. Grunwald, Über begrenzte Derivationen und deren Anwendung, Z. Angew. Math. Phys., 12 (1867) 441–480.
5. I. M. Gel'fand, G. E. Shilov, Generalized Functions, vol. 1, Academic Press, New York and London, 1964.
6. T. P. Higgins, The use of fractional integral operators for solving nonhomogeneous differential equations, Boeing Sci. Res. Laboratories Document, D1-82-0677, Dec., 1967.
7. J. Liouville, Mémoire: sur le calcul des différentielles à indices quelconques, J. de l'École Polytechnique, 13 (1832) 71–162.
8. T. P. G. Liverman, Generalized Functions and Direct Operational Methods, vol. 1, Prentice-Hall, Englewood Cliffs, N. J., 1964.
9. P. A. Nekrassov, General differentiation, (in Russian), Mat. Sb., 14 (1888) 45–168.
10. T. J. Osler, Leibniz rule for fractional derivatives generalized and application to infinite series, SIAM J. Appl. Math., 18 (1970) 658–674.
11. E. L. Post, Generalized differentiation, Trans. Amer. Math. Soc., 32 (1930) 723–781.
12. M. Riesz, L'intégral de Riemann-Liouville et le problème de Cauchy, Acta Math., 81 (1949) 1–233.
13. Y. Watanabe, Notes on the generalized derivative of Riemann-Liouville and its application to Leibniz's formula, I and II, Tôhoku Math. J., 34 (1931) 8–41.
14. E. T. Whittaker, G. N. Watson, A Course of Modern Analysis, 4th ed., Cambridge University Press, New York, 1962.
15. A. Zygmund, Trigonometric Series, Vol. II, Cambridge University Press, New York, 2nd ed., 1959.

ANOTHER SOLUTION OF AN OLD PROBLEM OF PÓLYA

SIMEON REICH, The Technion, Haifa, Israel

In 1913, Pólya [3] proposed the following problem: Show that there is no uniform affixing of $\pm$ signs to the elements of the square matrices of order $n>2$ such that the permanent of the resulting matrix equals the determinant of the original one. (The permanent of $A=\|a_{ij}\|$ is per $A=\sum a_{1\sigma(1)}a_{2\sigma(2)}\cdots a_{n\sigma(n)}$, where σ runs over all permutations of $\{1, \cdots, n\}$.) Szegö [4] solved this problem by using parity arguments. Meanwhile a much more general result was established in [2, p. 381], namely, there is no linear map $A\to T(A)$ on the square

matrices of order $n>2$ such that per $T(A)=\det A$ for all A. Nevertheless, it may be of interest to present the following simple solution of Pólya's problem:

First note that if $B=\|\pm 1\|$ is an n-square matrix, then det B is a multiple of 2^{n-1}. To see this, add the second column to the first one, expand from the new first column, and use induction. Now put

$$A = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}.$$

Were Pólya's assertion false, there would be a 3×3 matrix $B=\|\pm 1\|$ such that $6=\text{per } A=\det B=4k$, a contradiction. Finally, by considering $C=A\oplus I_{n-3}$, we obtain the desired conclusion for all $n\geqq 3$.

REMARK 1. Clearly for $n=2$ the required transformation exists.

REMARK 2. The contradiction could be reached by using Hadamard's inequality [1, p. 253]: $6=\text{per } A=\det B\leqq\sqrt{27}$.

References

1. F. R. Gantmacher, The Theory of Matrices, vol. 1, Chelsea, New York, 1960.
2. Marvin Marcus and Henryk Minc, On the relation between the determinant and the permanent, Illinois J. Math., 5 (1961) 376–381.
3. G. Pólya, Aufgabe 424, Archiv der Mathematik und Physik (3), 20 (1913) 271.
4. G. Szegö, Lösung zu 424, Archiv der Mathematik und Physik (3), 21 (1913) 291–292.

POLYNOMIALS WHICH COMMUTE WITH A TCHEBYCHEFF POLYNOMIAL

E. A. BERTRAM, University of Hawaii

Given a polynomial $Q(x)$ with coefficients in a field $\mathfrak{F}$ of characteristic 0, we may ask for all polynomials $P\in\mathfrak{F}[x]$ which are **permutable** or **commute** with Q, i.e., all $P(x)$ for which $P(Q(x))=Q(P(x))$ holds for all x. The polynomials x and $Q^{(k)}(x)$, the iterates of Q, are of course solutions, and if Q is itself the iterate of a third polynomial R, all iterates of R are solutions. In this note we characterize all solutions to this equation when $Q(x)$ is any polynomial from the sequence $\{T_n\}$ of Tchebycheff polynomials:

$$T_0 = 1, \qquad T_1(x) = x,$$
$$T_2(x) = 2x^2 - 1, \cdots, T_n(x) = 2xT_{n-1}(x) - T_{n-2}(x), \qquad n \geqq 2.$$

Since the techniques are valid for polynomials over any integral domain of characteristic 0, we shall work in this setting and prove that if $P(x)$ is a polynomial of degree $k\geqq 1$ which is permutable with some T_n, $n\geqq 2$, then $P=\pm T_k$ if n is odd, and $P=T_k$ if n is even.

We note that G. Julia [4] and J. F. Ritt [5], using more advanced topological and algebraic methods involving the theory of Riemann surfaces and monodromy groups, characterized those nonlinear complex polynomials P and Q which

are permutable: Define $R(z)$ and $S(z)$ to be **equivalent** via the polynomial

$$\lambda(z) = az + b, \qquad a \neq 0, \qquad \text{if } R(z) = \lambda^{-1}(S(\lambda(z))),$$

where

$$\lambda^{-1}(z) = \frac{1}{a}(z - b).$$

Then (i) if no iterate of one is identical with any iterate of the other, either P and Q are equivalent (via the same λ) to polynomials of the form az^n, or both are equivalent to polynomials of the form $\pm T_n(z)$; (ii) if P and Q have a common iterate, then there exists a polynomial G such that P and Q are equivalent to iterates of $\epsilon_i G$, where ϵ_1 and ϵ_2 are certain roots of unity.

More recently, E. Jacobsthal [3], using purely algebraic techniques valid for polynomials over any field of characteristic 0, proved the following related result: Suppose a set $\mathcal{S}$ of permutable polynomials contains at least one polynomial of each degree $\geqq 1$, and assume that $P(x)$ is a nonlinear polynomial which commutes with *each* polynomial in $\mathcal{S}$. Then $P(x)$ must belong to $\mathcal{S}$. Furthermore $\mathcal{S}$ contains exactly one polynomial of each degree and (see also [1], p. 332), up to equivalence, is either the sequence $\{x, x^2, x^3, \cdots\}$ or the sequence $\{T_n\}$. In [2] this result is proved with coefficients restricted to the real numbers but assuming the existence of (possibly complex) polynomial roots. These proofs that P must belong to $\mathcal{S}$ have all depended crucially on the assumption that P commutes with each polynomial in $\mathcal{S}$ and hence with a *quadratic* polynomial in $\mathcal{S}$. It is our object here to show that the elementary methods in [2] yield a characterization of those polynomials which commute with at least one T_n, $n \geqq 2$.

Each T_n has integer coefficients, and thus could be defined over any commutative ring with identity. Our proof assumes that all coefficients lie in an integral domain of characteristic 0; we use the facts that (i) the ring of polynomials over such a domain is also an integral domain, and (ii) if the polynomial $P(x) = \sum_{j=0}^{k} a_j x^j$ has the zero polynomial as its formal derivative $P'(x) = \sum_{j=1}^{k} j a_j x^{j-1}$, then $P(x)$ has degree 0. It should be pointed out that our result is valid over every integral domain of characteristic 0 if and only if it is valid over every field of characteristic 0.

Notation. Substitution of $Q(x)$ for the variable x in $P(x)$ is denoted either by $P(Q(x))$ or $P(Q)$. Ordinary polynomial multiplication is given by juxtaposition, as in $(1-x^2)T_{n-1}(x)$, or by the use of brackets, as in $A[P']^j$ and $(1-x^2)[T_n'(x)]^2$, to avoid the possibility of confusion with the substitution operation.

The first three lemmas can readily be proved by induction, appealing only to the recursive definition of T_n. On the other hand, considered as identities between polynomials with real integer coefficients, they are well-known identities over the real numbers and hence are valid over *any* integral domain of characteristic 0.

LEMMA 1. *$2T_jT_k = T_{j+k} + T_{j-k}$ for all $j \geqq k \geqq 0$.*

LEMMA 2. (Commutativity). $T_j(T_k) = T_k(T_j) = T_{jk}$.

LEMMA 3. *Each T_n satisfies the identity $(1-x^2)T_n'(x) + nxT_n(x) = nT_{n-1}(x)$.*

LEMMA 4. *Suppose $P(x)$ satisfies the following identity, for n a natural number:*

$$(1) \qquad (1 - x^2)[P'(x)]^2 = n^2[1 - P^2(x)].$$

Then $P(x)$ is either $\pm T_n(x)$.

Proof of Lemma 4. From Lemma 1 we have the four identities

$$T_{2n-2} + T_0 = 2T_{n-1}^2, \qquad 2x[T_{2n-1}(x) + T_1(x)] = 4xT_n(x)T_{n-1}(x),$$

$$T_{2n} + T_0 = 2T_n^2, \qquad \text{and} \qquad T_{2n-2} + T_{2n} = 2T_1T_{2n-1}.$$

After adding the first three of these and reducing using the fourth, we find that $T_{n-1}^2(x) - 2xT_n(x)T_{n-1}(x) + T_n^2(x) = 1 - x^2$. On the other hand, Lemma 3 shows that

$$(1 - x^2)^2[T_n'(x)]^2 = n^2(T_{n-1}^2(x) - 2xT_n(x)T_{n-1}(x) + T_n^2(x)) - n^2T_n^2(x) + n^2x^2T_n^2(x),$$

and the right side further reduces to $n^2(1-x^2)[1-T_n^2(x)]$. Since $1-x^2 \neq 0$, the identity $(1-x^2)[T_n'(x)]^2 = n^2[1-T_n^2(x)]$ follows.

To see that $\pm T_n(x)$ are the *only* polynomial solutions, formally differentiate both sides of (1). We obtain

$$(2) \qquad (1 - x^2)P''(x) - xP'(x) + n^2P(x) = 0,$$

since we may assume deg $P \geqq 1$. Let $P(x) = a_kx^k + a_{k-1}x^{k-1} + \cdots + a_0$, $k \geqq 1$ and $a_k \neq 0$. A comparison of the leading coefficients in (2) shows that $a_k[k(k-1) + k - n^2] = 0$, or $k = n$. If $t_n \neq 0$ is the coefficient of x^n in $T_n(x)$, the polynomial $t_nP(x) - a_nT_n(x)$ also satisfies (2) and has degree $< n$. This is impossible, unless $t_nP - a_nT_n$ is the zero polynomial. Now $P(1) = \pm 1$ follows from (1), and this together with $T_n(1) = 1$ gives $a_n = \pm t_n$ and $P(x) = \pm T_n(x)$.

THEOREM 1. *Suppose $A(x)$ is a polynomial of degree j, and $Q(x)$ is a polynomial of degree $n > 1$ which satisfies the identity*

$$(3) \qquad A(x)[Q'(x)]^j = n^jA(Q(x)).$$

Then, if P is a polynomial of degree k which is permutable with Q, P satisfies the same identity, with n replaced by k.

Proof. We introduce the polynomial $G = A[P']^j - k^jA(P)$, and shall show that G is the zero polynomial. If $G \neq 0$, it follows immediately that deg $G < kj$. We show that $G \neq 0$ also leads to deg $G = kj$. Using (3) and the commutativity, we obtain

$$n^jG(Q) = n^jA(Q)[P'(Q)]^j - k^jn^jA(P(Q)) = A[Q']^j[P'(Q)]^j - k^j[A(P)][Q'(P)]^j.$$

By the chain rule, $P'Q'(P) = Q'P'(Q)$, and we now have

$$n^jG(Q) = A[P']^j[Q'(P)]^j - k^jA(P)[Q'(P)]^j$$

$$= [Q'(P)]^j\{A[P']^j - k^jA(P)\} = [Q'(P)]^jG.$$

If $G \neq 0$, we compare degrees and find that $n \deg G = (n-1)kj + \deg G$, or $\deg G = kj$ since $n > 1$. The desired contradiction has been reached and the theorem proved.

THEOREM 2. *Let $\{T_n\}_{n \geq 2}$ be the sequence of nonlinear Tchebycheff polynomials, and P a polynomial of degree $k \geq 1$ which is permutable with at least one T_n. Then $P = T_k$ if n is even, and $P = \pm T_k$ if n is odd.*

Proof. By Lemma 4, $\pm T_n$ are the only polynomials P of degree n which satisfy the identity $A[P']^2 = n^2A(P)$, where $A(x)$ is the polynomial $1-x^2$. But the hypothesis that $T_n(P) = P(T_n)$, $n > 1$, and Theorem 1 imply that P must satisfy this identity, with n replaced by k. Thus $P = \pm T_k$. If n is odd, T_n is an odd function and $P = \pm T_k$; if n is even, T_n is an even function, and $P = -T_k$ is impossible.

References

1. T. Bang, Congruence properties of Tchebycheff polynomials, Math. Scand., 2 (1954) 327–333.

2. H. D. Block and H. P. Thielman, Commutative polynomials, Quart. J. Math. Oxford Ser., (2), 2 (1951) 241–3.

3. E. J. Jacobsthal, Über vertauschbare Polynome, Math. Z., 63 (1955) 243–276.

4. G. Julia, Mémoire sur la permutabilité des fractions rationnelles, Ann. Sci. École Norm. Sup., 39 (1922) 131–215.

5. J. F. Ritt, Permutable rational functions, Trans. Amer. Math. Soc., 25 (1923) 399–448.

A NOTE ON IDEMPOTENTS IN BANACH ALGEBRAS

A. R. BLASS, University of Michigan, AND C. V. STANOJEVIC, University of Missouri at Rolla

In [1, p. 284] the following theorem is proved:

THEOREM. *Suppose B is a real or complex Banach algebra, and $f:(0, \infty)\to B$ is a function which is measurable and (B)-integrable in each finite interval $(0, \omega)$ and which satisfies*

$$f(\xi_1 + \xi_2) = f(\xi_1)f(\xi_2); \qquad \xi_1, \xi_2 \in (0, \infty).$$

If

$$\lim_{\xi\to 0^+}\frac{1}{\xi}\int_0^{\xi} f(\tau)d\tau = j$$

exists, then j is an idempotent.

In this note we give the following version of the above theorem:

THEOREM. *Let B be a real or complex Banach algebra with unit e. A necessary and sufficient condition that B has a nontrivial idempotent element b is that there exists a nonconstant mapping $f\colon(0,\infty)\to B$ such that:*

(1) $f(\xi_1\xi_2) = f(\xi_1)f(\xi_2)$; $\quad \xi_1, \xi_2 \in (0, \infty)$.

(2) *f is (B)-integrable on each finite* $(0, \omega)$.

(3) $\displaystyle\lim_{\eta\to 0^+}\frac{1}{\eta}\int_0^{\eta} f(\tau)d\tau = b$ *exists and* $b \neq \theta$.

Proof. Necessity. Suppose $b\neq\theta$, e and $b^2=b\in B$. Define $f\colon(0,\infty)\to B$ by

$$f(\xi) = \xi e + (1-\xi)b.$$

Then (1) follows by direct calculation. Since

$$\int_0^{\xi} f(\tau)d\tau = \frac{\xi^2}{2}e + \left(\xi - \frac{\xi^2}{2}\right)b,$$

(2) holds. Also,

$$\lim_{\eta\to 0^+}\frac{1}{\eta}\int_0^{\eta} f(\tau)d\tau = \lim_{\eta\to 0^+}\left(\frac{\eta}{2}e + \left(1-\frac{\eta}{2}\right)b\right) = b;$$

thus (3) holds.

Sufficiency. Consider $(1/\eta)\int_0^\eta f(\tau)d\tau$, where f satisfies (1), (2), and (3). First, due to (1), we have

$$f(\xi)\frac{1}{\eta}\int_0^{\eta} f(\tau)d\tau = \frac{1}{\eta}\int_0^{\eta} f(\xi)f(\tau)d\tau = \frac{1}{\eta}\int_0^{\eta} f(\xi\tau)d\tau.$$

Changing the variable to $\gamma=\xi\tau$ on the right, we obtain

$$f(\xi)\frac{1}{\eta}\int_0^{\eta} f(\tau)d\tau = \frac{1}{\xi\eta}\int_0^{\xi\eta} f(\gamma)d\gamma.$$

By (3), we obtain

$$(*) \qquad f(\xi)b = b; \qquad \xi > 0.$$

Hence, for $\xi>0$, from (*) we have $(1/\xi)\int_0^\xi f(\tau)d\tau\cdot b = (1/\xi)\int_0^\xi b\,d\tau = b$. Let $\xi\to 0^+$; it follows that $b^2=b$. Note that $b\neq e$, since otherwise it would follow from (*) that f is constant. Since $b\neq\theta$ by hypothesis, this completes the proof.

Reference

1. E. Hille and R. Phillips, Functional Analysis and Semigroups, AMS Colloq. Publ., 31 (1957).

ISOMETRIES IN NORMED SPACES

J. A. Baker, University of Waterloo

Mazur and Ulam [5] have shown that an isometry f of one real normed linear space *onto* another is necessarily affine (i.e., $x \to f(x) - f(0)$ is linear). It is natural to ask if the result holds without the onto assumption. In this note we prove that an isometry from one real normed linear space into a *strictly convex* real normed linear space is affine. We also show that, given a real normed linear space Y which is not strictly convex, there exists an isometry from the reals into Y which is not affine.

For further discussions concerning the result of Mazur and Ulam, see Day [3], page 110, and Charzyński [1].

By an **isometry** from a normed linear space X into a normed linear space Y we mean a function $f: X \to Y$ such that $\|f(x) - f(y)\| = \|x - y\|$ for all $x, y \in X$. We say that a normed linear space Y is **strictly convex** provided $a, b \in Y$ and $\|a+b\| = \|a\| + \|b\|$ imply $\{a, b\}$ is linearly dependent. It is easy to see that a normed linear space Y is strictly convex if and only if $a, b \in Y$, $a \neq 0$, $b \neq 0$, and $\|a+b\| = \|a\| + \|b\|$ imply $a = tb$ for some $t > 0$. The following result appears in Fischer and Muszély [4]:

Lemma 1. *If Y is a normed linear space, $a, b \in Y$, and $\|a+b\| = \|a\| + \|b\|$, then $\|sa+tb\| = s\|a\| + t\|b\|$ for all $s, t \geqq 0$.*

Proof. Suppose without loss of generality that $0 \leqq s \leqq t$. Then

$$\|sa + tb\| \leqq s\|a\| + t\|b\|$$

and, on the other hand,

$$\begin{aligned} \|sa + tb\| &= \|t(a + b) - (t - s)a\| \\ &\geqq |\, t\|a + b\| - (t - s)\|a\|\, | \\ &= s\|a\| + t\|b\|. \end{aligned}$$

Lemma 2. *Let Y be a real normed linear space which is strictly convex and let $a, b \in Y$. Then $\frac{1}{2}(a+b)$ is the unique member of Y, which is a distance $\frac{1}{2}\|a-b\|$, from both a and b.*

Proof. The result clearly holds if $a = b$. It is also easy to see that $\frac{1}{2}(a+b)$ is a distance $\frac{1}{2}\|a-b\|$ from both a and b. Thus it suffices to prove the uniqueness.

Suppose then that $a \neq b$ and $u, v \in Y$ with

$$\begin{aligned} \|a - u\| &= \|b - u\| = \|a - v\| = \|b - v\| \\ &= \tfrac{1}{2}\|a - b\|. \end{aligned}$$

Then

$$\begin{aligned}(1)\qquad \|a - \tfrac{1}{2}(u+v)\| &= \|\tfrac{1}{2}(a-u) + \tfrac{1}{2}(a-v)\| \\ &\leqq \tfrac{1}{2}\|a-u\| + \tfrac{1}{2}\|a-v\| = \tfrac{1}{2}\|a-b\|.\end{aligned}$$

Similarly,

$$(2)\qquad \|b - \tfrac{1}{2}(u+v)\| \leqq \tfrac{1}{2}\|a-b\|.$$

If either of these inequalities were strict we would have

$$\begin{aligned}\|a-b\| &\leqq \|a - \tfrac{1}{2}(u+v)\| + \|b - \tfrac{1}{2}(u+v)\| \\ &< \|a-b\|.\end{aligned}$$

Thus equality holds in (1) and (2) so that

$$\begin{aligned}\|\tfrac{1}{2}(a-u) + \tfrac{1}{2}(a-v)\| &= \tfrac{1}{2}\|a-b\| \\ &= \|\tfrac{1}{2}(a-u)\| + \|\tfrac{1}{2}(a-v)\|.\end{aligned}$$

Since Y is strictly convex, $a \neq u$ and $a \neq v$, it follows that $a-u=t(a-v)$ for some $t>0$. But, since $\|a-u\|=\|a-v\|$, $t=1$ and thus $u=v$.

We can now prove the main result of this note.

THEOREM. *Let X and Y be real normed linear spaces and suppose Y is strictly convex. If $f: X \to Y$ is such that*

$$(3)\qquad \|f(x) - f(y)\| = \|x-y\|$$

for all $x, y \in X$, then f is affine.

Proof. If $f(0) \neq 0$, let $g(x) = f(x) - f(0)$. Then g is an isometry and $g(0)=0$. Thus we may assume that $f(0)=0$.

Putting $y=0$ in (3) we find $\|f(x)\| = \|x\|$, and hence $\|f(-x)\| = \|x\|$ for all $x \in X$. Replacing y by $-x$ in (3) we obtain

$$\|f(x) + (-f(-x))\| = 2\|x\| = \|f(x)\| + \|-f(-x)\|.$$

Since Y is strictly convex we must have $f(x) = -tf(-x)$ for some $t>0$. But $\|f(x)\| = \|f(-x)\|$ so $t=1$. Thus $f(-x) = -f(x)$ for all $x \in X$.

Now, for all $x, y \in X$,

$$\begin{aligned}\|f(x+y)\| &= \|f(x+y) - f(0)\| = \|x+y\| \\ &= \|x-(-y)\| = \|f(x) - f(-y)\| \\ &= \|f(x) + f(y)\|.\end{aligned}$$

Hence

$$\begin{aligned}\left\|f\left(\frac{x+y}{2}\right) - f(x)\right\| &= \left\|\frac{x-y}{2}\right\| = \frac{1}{2}\|x-y\| \\ &= \frac{1}{2}\|f(x) - f(y)\|\end{aligned}$$

and similarly,

$$\left\| f\left(\frac{x+y}{2}\right) - f(y) \right\| = \frac{1}{2} \|f(x) - f(y)\|$$

for all $x, y \in X$. It follows from Lemma 2 that

$$f\left(\frac{x+y}{2}\right) = \frac{f(x)+f(y)}{2} \tag{4}$$

for all $x, y \in X$. Since $f(0)=0$ we easily find from (4) that f is additive. But, being an isometry, f is continuous. Hence f is linear and this completes the proof.

Let Y be a real normed linear space which is not strictly convex. Following [4] we give an example of an isometry from $\mathfrak{R}$, the reals, into Y which is not affine. To this end, choose $a, b \in Y$ such that $\{a, b\}$ is linearly independent, $\|a\| = \|b\| = 1$, and $\|a+b\| = \|a\| + \|b\|$. This is possible according to Lemma 1 and the definition of strict convexity. For $x \in \mathfrak{R}$ define

$$f(x) = \begin{cases} xa & \text{if } x \leqq 1 \\ a + (x-1)b & \text{if } x > 1. \end{cases}$$

It is easy to verify, using Lemma 1, that f is an isometry but is not affine.

As a further example, let us construct a *homogeneous* isometry which is not linear. This answers a problem raised in [2].

Let $g: \mathfrak{R}^2 \to \mathfrak{R}$ be defined by

$$g(x, y) = \begin{cases} y & \text{if } 0 \leqq y \leqq x \quad \text{or} \quad x \leqq y \leqq 0 \\ x & \text{if } 0 \leqq x \leqq y \quad \text{or} \quad y \leqq x \leqq 0 \\ 0 & \text{otherwise.} \end{cases}$$

It is not difficult to see that

(i) g is homogeneous, i.e., $g(tx, ty) = tg(x, y)$ for all $t, x, y \in \mathfrak{R}$,

(ii) $|g(x, y) - g(u, v)| \leqq \sqrt{(x-u)^2 + (y-v)^2}$ for all $x, y, u, v \in \mathfrak{R}$,

(iii) g is not linear.

Let X denote $\mathfrak{R}^2$ with the usual normed linear space structure, and let Y denote $\mathfrak{R}^3$ with the usual vector space structure but with

$$\|(x, y, z)\| = \max(\sqrt{x^2 + y^2}, |z|).$$

Then, with this norm, Y is a normed linear space.

If $f: X \to Y$ is defined by $f(x, y) = (x, y, g(x, y))$, then it follows from (i), (ii), and (iii) that f is a homogeneous isometry which is not linear.

This work has been supported by N.R.C. Grant A7153.

References

1. Z. Charzyński, Sur les transformations isométriques des espaces du type (F), Studia Math., 13 (1953) 94–121.

2. P. R. Chernoff, Advanced Problem 5688, this MONTHLY, 76 (1969) 835.
3. M. M. Day, Normed Linear Spaces, Academic Press, New York, 1962.
4. P. Fischer and Gy. Muszély, On some new generalizations of the functional equation of Cauchy, Canad. Math. Bull., 10 (1967) 197–205.
5. S. Mazur, and S. Ulam, Sur les transformations isométriques d'espaces vectoriels normes, C. R. Acad. Sci. Paris, 194 (1932) 946–948.

A MULTINOMIAL SUM

J. B. KADANE, Center for Naval Analyses and Carnegie-Mellon University

The following sum arose from a statistical problem [1]:

$$\sum \prod_{i=1}^{m} \prod_{j=1}^{n} \frac{1}{(x_{ij})!} = \frac{N!}{\prod_{i=1}^{m} a_i! \prod_{j=1}^{n} b_j!}, \tag{1}$$

where the summation is over the set S of all nonnegative integers x_{ij} $(i=1, \cdots, m; j=1, \cdots, n)$ satisfying $\sum_{j=1}^{n} x_{ij}=a_i$ for each $i=1, \cdots, m$, $\sum_{i=1}^{m} x_{ij}=b_j$ for each $j=1, \cdots, n$; for consistency we assume

$$\sum_{i=1}^{m} a_i = \sum_{j=1}^{n} b_j = N.$$

Formula (1) can be expressed in an asymmetric but more familiar form: Let

$$(n_1, n_2, \cdots, n_r) = \frac{(n_1 + n_2 + \cdots + n_r)!}{n_1!\, n_2! \cdots n_r!}.$$

Then (1) is equivalent to

$$\sum \prod_{i=1}^{m} (x_{i1}, \cdots, x_{in}) = (b_1, \cdots, b_n), \tag{2}$$

where again the summation is over S. Viewed as (2), (1) is a generalization of the multivariate Vandermonde equality [2, 3, 4].

Proof: Equate coefficients in the expansion of

$$\prod_{i=1}^{m} (y_1 + \cdots + y_{n-1} + 1)^{a_i} = (y_1 + \cdots + y_{n-1} + 1)^N,$$

which yields (2).

After writing this note, I found (2), with a more complicated proof, in [5]. I should like to thank John Riordan for his bibliographic help.

References

1. G. R. Iversen, J. B. Kadane, and W. F. Rogers, Estimation of multinomial probabilities when only the sum is observed (in preparation).

2. L. Carlitz, Sums of products of multinomial coefficients, Elem. Math., 18 (1963) 37–39.
3. J. Riordan, Combinatorial Identities, Wiley, New York, 1968, p. 38.
4. S. Tauber, On multinomial coefficients, this MONTHLY, 70 (1963) 1058–1063.
5. S. S. Wilks, Mathematical Statistics, Princeton University Press, Princeton, N. J., 1943, p. 215, eq.(h).

RESEARCH PROBLEMS

EDITED BY RICHARD GUY

In this Department the Monthly presents easily stated research problems dealing with notions ordinarily encountered in undergraduate mathematics. Each problem should be accompanied by relevant references (if any are known to the author) and by a brief description of known partial results. Manuscripts should be sent to Richard Guy, Department of Mathematics, Statistics, and Computing Science, The University of Calgary, Calgary 44, Alberta, Canada.

LINEAR TRANSFORMATIONS OF A FINITE FIELD

S. E. PAYNE, Miami University

Let F be a finite field with 2^e elements. Then the group of automorphisms of F is cyclic of order e with generator ρ: $x \to x^2$, $x \in F$. For each i, $1 \leqq i \leqq e$, with g.c.d. $(i, e) = 1$, $\alpha = \rho^i$ has the following two properties:

(1) α is a nonsingular linear transformation of F over the prime subfield, i.e.,

$$(x + y)^\alpha = x^\alpha + y^\alpha, \quad x, y \in F.$$

(2) The map T_α defined by

$$T_\alpha\colon x \to \begin{cases} 0, & x = 0 \\ x^\alpha x^{-1}, & x \neq 0 \end{cases}$$

is a permutation of the elements of F.

If α is any additive map of F satisfying (1) and (2), then α_c defined by $\alpha_c\colon x \to cx^\alpha$ for $0 \neq c \in F$, $x \in F$, also satisfies (1) and (2). The problem we have in mind is to determine whether or not the preceding discussion determines **all** the linear maps satisfying (1) and (2).

PROBLEM. Determine all nonsingular additive maps α of F such that $x \to x^\alpha x^{-1}$ permutes the nonzero elements of F (and, without loss of generality, $1^\alpha = 1$).

The interest in such α lies in the fact that they yield ovoids in the plane coordinatized by F (cf., [2], p. 50), which in turn yield generalized quadrangles (cf. [4]). Indeed, the more general problem, which is equivalent to asking for a complete determination of all ovoids in the projective plane $PG(2, 2^e)$, is stated implicitly in [4] as follows:

GENERAL PROBLEM. Determine all permutations α of F satisfying $0^\alpha = 0$, $1^\alpha = 1$, and

$$\frac{c_0 - c_1}{c_0^\alpha - c_1^\alpha} \neq \frac{c_0 - c_2}{c_0^\alpha - c_2^\alpha} \tag{3}$$

for all triples c_0, c_1, c_2 of distinct elements of F.

Some preliminary remarks on the General Problem may be found in [5], but any constructions of such α which are linear and not automorphisms would seem to be new. We are indebted to the referee for pointing out that in the linear case we may suppose α to be given by

$$(4) \qquad x^\alpha \equiv \sum_{i=0}^{e-1} a_i x^{2^i},\ x^\alpha x^{-1} = \sum_{i=1}^{e-1} a_i\, x^{2^i-1}, \qquad \text{for fixed } a_i \in F,$$

and for each $x \in F$.

The nonsingularity part of (1) is then automatically satisfied if (2) is satisfied; the problem is essentially one of finding a polynomial of the type given for $x^\alpha x^{-1}$ which induces a permutation on the nonzero elements of F.

Carlitz [1] and McConnel [3] have solved problems which seem related somewhat to the General Problem. However, we see no way to apply their results to gain information about the problem under consideration here.

References

1. L. Carlitz, A theorem on permutations in a finite field, Proc. Amer. Math. Soc., 11 (1960) 456–459.
2. P. Dembowski, Finite Geometries, Springer-Verlag, New York, 1968.
3. R. McConnel, Pseudo-ordered polynomials over a finite field, Acta Arith., 8 (1963) 127–151.
4. S. E. Payne, Affine representations of generalized quadrangles, J. Algebra, 16 (1970) 473–485.
5. ———, Generalized quadrangles as amalgamations of projective planes (unpublished).

CLASSROOM NOTES

EDITED BY DAVID DRASIN

Manuscripts for this Department should be sent to Robert Gilmer, Department of Mathematics, Florida State University, Tallahassee, FL 32306.

THE LEBESGUE DECOMPOSITION THEOREM FOR MEASURES

J. K. BROOKS, University of Florida

In this note we present a short and simple proof of the Lebesgue decomposition theorem for measures. The method of proof enables us to decompose a measure uniquely into a continuous and a singular part with respect to an outer measure. Most of the standard proofs (see [3] or [4]) use the Radon-Nikodým theorem, which is a deep result, and which in the case of outer meaures, is not applicable. In [5] a proof is given that avoids the use of the Radon-Nikodým theorem, but it is long and technically difficult. For decomposition theorems for measures whose range lies in a Banach space or the hyperspace of a Banach space, the reader is referred to [1] or [2] respectively.

Let Σ denote a σ-algebra of subsets of a set S. A set function defined on Σ

2. J. K. Brooks, An integration theory for set-valued measures, I, Bull. Soc. Roy. Sci. Liège, 5–8 (1968) 312–319.
3. P. R. Halmos, Measure Theory, Van Nostrand, New York, 1950.
4. E. Hewitt and K. Stromberg, Real and Abstract Analysis, Springer-Verlag, New York, 1965.
5. S. Saks, Theory of the Integral, 2nd ed., vol. 7, Monografje Matematyczne, Warsaw, 1937.

MAPPINGS WITH NON-VANISHING JACOBIAN

H. KESTELMAN, University College, London

1. A mapping of R^2 into itself can have a non-vanishing continuous Jacobian without being injective: for example, the transformation of the Argand plane by e^z. In this note it is proved that the mapping must be injective if certain boundary conditions are satisfied. We are interested mainly in differentiable mappings f of an open subset G of R^m into R^m when Df is continuous and invertible at all points of a compact subset K of G; but the essentials of the proofs become clearer in a more general situation.

Let V be a normed vector space over the reals and G an open subset of V which is mapped continuously into V by a function F with the following properties: (a) F transforms every open subset of G into an open set, (b) F is *locally injective* at all points of G, i.e., to every ξ in G there is a neighborhood N of ξ such that F_N is injective (we write F_E for the restriction of F to any subset E of G). We say that x in E is a *multiple point* of F_E if $F(x)=F(y)$ for some y in E other than x; the set of all such x will be denoted by M_E. The results to be proved are:

THEOREM 1. *If K is a compact subset of G and there are no points of M_K in ∂K, then F_K is injective.*

THEOREM 2. *If K is a compact subset of G, and ∂K is connected, and $F_{\partial K}$ is injective, then F_K is injective (∂E denotes the boundary of E, and E° will denote the interior of E).*

The assumption in Theorem 2 that ∂K is connected cannot be omitted: thus, the mapping of the union of the closed discs $|z|\leqq 1$ and $|z-2\pi i|\leqq \frac{1}{2}$ by the function e^z is not injective even though its restriction to the boundary is so.

2. LEMMA 1. *Let K be a compact subset of V, and H a continuous mapping of K into V which is locally injective at all points of K; then the set of multiple points of H is closed.*

Proof. Let $x_1, x_2, \cdots$ be multiple points of H converging to ξ (in K); we have to show that $H(\xi)=H(\eta)$ for some η in K other than ξ. There is a sequence $\{y_n\}$ in K with $y_n\neq x_n$ and $H(y_n)=H(x_n)$; K being compact, we may assume that $\{y_n\}$ converges to some η in K. Since H is continuous, $H(\xi)=H(\eta)$; to

show that $\xi \neq \eta$ we observe that ξ is the center of a ball B in which H is injective, so that y_n is outside B for all large n and hence $\xi \neq \eta$.

3. *Proof of Theorem* 1. We assume that F satisfies (a) and (b) of Section 1, that K is a compact subset of G, that $M_K \cap \partial K = \emptyset$, and that F_K is not injective; we then deduce a contradiction. Since M_K is compact by Lemma 1, its distance from the nonempty compact set ∂K is equal to $\|\alpha - \beta\|$ for some α in M_K and some β in ∂K. Choose α^* in K so that $\alpha^* \neq \alpha$ and $F(\alpha^*) = F(\alpha)$; then $\alpha^* \in K^\circ$ since $M_K \cap \partial K = \emptyset$. Let B be a ball in K with center α^* and radius less than $\|\alpha - \alpha^*\|$. By (a), $F(\alpha)$ (equal to $F(\alpha^*)$) is interior to $F(B)$. Since $\alpha \in K^\circ$, all x on the segment (α, β) which are close enough to α belong to $K \backslash B$ while $F(x)$, like $F(\alpha)$, belongs to $F(B)$; hence such x are in M_K and $\|x - \beta\| < \|\alpha - \beta\|$, contradicting the definition of α and β.

4. *Proof of Theorem* 2. We assume that F satisfies (a) and (b) of Section 1, that K is compact, with ∂K connected, and that $F_{\partial K}$ is injective while F_K is not; from this we deduce a contradiction. By Theorem 1, $M_K \cap \partial K$ is nonempty and is closed (Lemma 1); to reach a contradiction we prove that $\partial K \backslash M_K$ is nonempty and closed. $\|F\|$ being continuous on K, ξ exists in K with $\|F(\xi)\| \geqq \|F(x)\|$ for all x in K. By (a), every α in K° is the center of a ball B with $F(\alpha)$ interior to $F(B)$ and hence $\|F(x)\| > \|F(\alpha)\|$ for some x in B. Hence ξ is not in K°, i.e., $\xi \in \partial K$, and $F(\alpha) \neq F(\xi)$ for all α in K°; at the same time $F(x) \neq F(\xi)$ for all x in ∂K with $x \neq \xi$. It follows that ξ is in ∂K but not in M_K.

To prove that $\partial K \backslash M_K$ is closed, suppose if possible that this is false and that ζ is a point in M_K which is the limit of a sequence $\{x_n\}$ in $\partial K \backslash M_K$. Choose η in K so that $F(\eta) = F(\zeta)$ and $\eta \neq \zeta$. Since $F_{\partial K}$ is injective and $\zeta \in \partial K$, $\eta \in K^\circ$. By (a), since $F(\zeta) = F(\eta)$, η is the center of a ball B in K° with $F(\zeta)$ interior to $F(B)$, and so $F(x_n) \in F(B)$ for all large n. If we make the radius of B smaller than $\|\eta - \zeta\|$, we reach the contradiction that $x_n \in M_K$ for some n.

5. Returning to our original problem concerning mappings of R^m into itself, we now have:

THEOREM 3. *Suppose* (i) *f is a differentiable mapping of an open subset G of R^m into R^m,* (ii) *Df is continuous and invertible at all points of K, a compact subset of G. Then f_K is injective if* **either** *of the following conditions is satisfied*:

(i) *there are no multiple points of f_K on ∂K,*

(ii) *$f_{\partial K}$ is injective and ∂K is connected.*

Proof. It is proved in advanced calculus that (i) and (ii) imply that (a) and (b) of Section 1 hold when $F = f$. The theorem is now a consequence of Theorems 1 and 2.

MATHEMATICAL EDUCATION

EDITED BY J. G. HARVEY AND M. W. POWNALL

Material for this Department should be sent to either of the editors: J. G. Harvey, Department of Mathematics, University of Wisconsin, Madison, WI 53706; M. W. Pownall, Department of Mathematics, Colgate University, Hamilton, NY 13346.

CALCULUS AS AN EXPERIMENTAL SCIENCE*

R. P. BOAS, JR., Northwestern University

I hope that my title was not too misleading. I am not going to suggest that calculus should somehow be based on experiment, but rather that calculus should be presented to the student in the same spirit as the experimental sciences. The point that I hope to make is, briefly, that proofs are to mathematics what experiments are to physics (or chemistry, or biology), and that our teaching can profit by the analogy.

Let us first of all be clear about what calculus is. There are two big ideas, the derivative and the integral. Geometrically, these are the slope of a curve and the area under a curve. Of course they frequently, even usually, appear in non-geometrical forms: a derivative might represent a mass density and an integral might represent work, for instance. However, translating to and from geometry should not bother anyone who has ever done something like drawing a vector diagram of forces. The predecessors of Newton and Leibnitz knew perfectly well how to determine tangents and areas, but they had to approach each problem from first principles. The great contribution of Newton and Leibnitz was precisely to make the procedures for finding tangents, areas, etc. into a calculus, that is, a systematic way of calculating—a collection of algorithms, to use the currently fashionable word. Moreover, they didn't really understand—in the modern sense—why the algorithms worked. Perhaps it will make the point clearer if I use a very elementary example. If you write two numbers with Roman numerals, and want to multiply them, you can work out the product if you understand the commutative, associative, and distributive laws of multiplication and addition for integers, but it takes time. Presumably the Romans used some other method, perhaps some kind of abacus (a simple digital computer). A more convenient way is to use Arabic numerals and the rules for manipulating them—a kind of calculus, in fact. In either case we have substituted rules or

* Vice-presidential address given at the meeting of the American Association for the Advancement of Science, December 28, 1970, as part of a Symposium on Mathematics in the Undergraduate Program in the Sciences, jointly sponsored by CUPM. The opinions expressed are those of the author, and are not to be taken as representing CUPM policy.

This article is also appearing in the Two Year College Mathematics Journal. Since Professor Boas is a prominent analyst and has recently served as chairman of the Committee on the Undergraduate Program in Mathematics, the editors feel that the mathematical community at large would be especially interested to know what he has to say about the teaching of elementary calculus. We have therefore departed from normal procedures, and duplicated publication of this article. *The Editors.*

procedures for thinking about what is actually going on. Note incidentally that one can often use a calculus successfully without fully understanding why it works. (Does a digital computer understand arithmetic?)

Once invented, differential and integral calculus were very successful at solving certain kinds of geometrical problems, and hence physical problems that can be represented geometrically, and hence problems in physics, chemistry, economics, etc., even when they are not represented geometrically. Consequently calculus became the standard language for talking about the subjects in which it was most successful, it has remained so to this day, and seems likely to continue for some time into the future. This is why every scientist has to study calculus, although he often wonders at first why he should have to. Another way of saying much the same thing is that calculus is used because it facilitates the study of models of observed phenomena. If a biological process, for instance, can be modeled as a differential equation, calculus can take over and predict properties of the process without using any biological thought, and the biologist can then compare the prediction with experiment—in this way he may save considerable time and thought. It is hard to get this idea across to the beginning student, especially when he doesn't know any biology yet.

Everybody admits, I suppose, that the sciences other than mathematics are based on experiment. Things that can be checked by experiment are accepted: things that disagree with experiment are not. However, I am not aware of any physics or chemistry or biology course that repeats all the classical experiments, or even any of those that are particularly difficult or time-consuming. At least in the science courses I took (of course, this was a long time ago) we were told that certain things had been established experimentally, and maybe (not always) what the experiment was like. Inspection of some current textbooks suggests that things haven't changed much in 40 years.

Now I do not know any experimental scientists who seem to feel uncomfortable about this state of affairs, although for all I know they may worry about it in secret. Nor do they seem to worry about the necessity of sometimes giving oversimplified or even mildly fallacious reasons why the experiment comes out as it does. For example, why does an airplane stay up? Elementary texts give theoretical reasons that do not seem very convincing; the real theoretical reasons are clearly too sophisticated for elementary courses; it presumably would be possible also to rely on experimental measurements of the flow around a wing, but few, if any, physics courses bring wind-tunnels into the classroom. Similarly, first-year physics courses usually teach Newtonian mechanics, rather than relativistic mechanics or quantum mechanics. In calling attention to this, I do not intend to criticize the current teaching of the experimental sciences; in fact, I do not see what else could be done, and indeed I want to use the experimental scientists' approach as a model.

Mathematics is not at all an experimental science, but there is a rather exact parallel between mathematics and the experimental sciences. In mathematics we believe things, not because we did an experiment, but because we proved

them. At least—and this continues the analogy—we believe things because *somebody* proved them; we have not necessarily studied the proof ourselves. Thus the proof is to mathematics as the experiment is to physics, chemistry, biology, and so on. Perhaps we are better off than the experimental scientist in one respect—we know that we *could* read the proof if we tried, whereas the experiment may be too difficult or too expensive for the experimental scientist to hope to repeat for himself or his class. Of course we want our students to believe what we say because we have done something that really carries conviction. What shall we do? The answer that most mathematicians believe in, or profess to believe in, or act as if they believed in, is that they ought to present their students with formal proofs of everything that they tell them. The effect of this is to make calculus a chapter in the theory of functions of a real variable. There are several reasons for this attitude. There are mathematicians who didn't understand calculus themselves to begin with, but now do; and, filled with missionary zeal, wish to spread the light. There are those who feel that it is intellectually dishonest not to tell everything that they know. There are those who feel that anything less than full explanations cheats the student. And there are those who understand the proofs but can't solve the problems: theory is always easier than technique. In any case, only just so much time is available. In order to make the best use of it, I claim that the teacher of calculus would do well to follow the lead of the experimental scientist: let him give proofs when they are easy and justify unexpected things; let him omit tedious or difficult proofs, especially those of plausible things. Let him give easy proofs under simplified assumptions rather than complicated proofs under general hypotheses. Let him by all means always give correct statements, but not necessarily the most general ones that he knows.

Let us see how these principles apply to some topics in calculus. (1) One of the unexpected results is the formula for the derivative of a product. Most beginners will guess it wrong. The proof is easy and completely convincing. One should by all means give it. (2) A function with a positive derivative is increasing. This looks, but isn't, tautological; the point at issue, generating a global property from a local one, is rather subtle. The proof is not illuminating, and might well be skipped. (3) It is certainly necessary to define the definite integral, but to prove that the integral of a continuous function exists is both technically demanding and time-consuming. This seems to be a clear case for "it can be proved." (4) The uniqueness theorem for solutions of a second-order linear differential equation is only too plausible—don't the initial position and velocity determine the motion? The proof is time-consuming, but the facts are easy to state precisely and meaningfully. (5) Assuming that Fourier series get into the calculus (they usually don't), it would be difficult, time-consuming, and unconvincing to *prove* any really useful convergence theorem. On the other hand, should Fourier series be left out just because we cannot prove a satisfying theorem about them? It is easy enough to state one, and there is no excuse for stating an incorrect one.

I am going to be accused by my colleagues of advocating a cookbook approach to calculus. This I deny. There was once a really cookbook approach to calculus, in which the student had to listen to incomprehensible nonsense until he developed a sound intuition (if he ever did). The approach that some of my colleagues favor makes the student listen to incomprehensible sense instead. I think the experimental scientists do better. Let me illustrate the difference with an example. A cookbook approach to maximum and minimum problems leads the student to approach all problems by setting a derivative equal to zero and testing by the sign of the second derivative. This traditional procedure can, in fact, lead to mistakes. A rigorous approach demands a long series of preliminary theorems about maxima, mean values and derivatives. What I prefer is something like this: observe that if there is a maximum where there is a derivative, the derivative must be zero; then the maximum must be at one of the (usually small number of) points where the derivative is zero or doesn't exist, or else at an endpoint. A small amount of computation will usually decide; and we avoid the second-derivative test, which in spite of its theoretical elegance is usually quite impractical. It seems to me that an approach of this kind is very much in the spirit in which experimental sciences are usually presented; and in practice it seems to give the students more capability with calculus, and sooner, than the theorem-proving approach that has been so popular.

WHAT RESEARCH COMPETENCIES FOR THE MATHEMATICS EDUCATOR?

H. H. WALBESSER, University of Maryland, and THEODORE EISENBERG, Northern Michigan University

Introduction. This article is divided into two sections. The first section contains a discussion of why there is a need for researchers in mathematics education. A rationale is presented for incorporating research competencies into doctoral programs. The proposed purpose of the research competencies is to enable one to query into the psychological and methodological dimensions of mathematics teaching and learning. The mathematics competencies associated with the bidisciplinary degree, mathematics education, are not discussed, because ample attention has been paid to them in previous publications [**2**, **3**, **4**, **6**]. The second section of this article outlines a sequence of seminars given at the University of Maryland. The seminars are designed to assist the student in the acquisition of specific research competencies for attacking problems in mathematics education. One seminar is described in detail, on task analysis competencies; its purpose is to acquaint the student to a behavioral approach to research on instruction. Brief descriptions of other seminars in the sequence are also provided.

1. **Training for Research—Why?** The university is an institution where the accumulated knowledge of man is exposed to public view for critical examination. But the university is more than a repository and retrieval center of past

accomplishments; it is also society's principal training ground for those who are to be the future knowledge producers. Doctoral programs are especially conceived to fulfill this training responsibility. Any doctoral program in mathematics education must hold the expectation of producing knowledge makers as well as knowledge consumers. Many of the present programs in mathematics education do not produce active, contributing researchers. The rarity of research among the total published literature of mathematics education attests to this failure.

The published research of a discipline is a measure of its contribution to man's knowledge of himself, his origins, and his environment. The acknowledgement of an individual as a scholar in any discipline does, therefore, reside with his research contributions and not with a demonstration that a certain collection of course requirements has been satisfied. The initial demonstration that one has acquired the competencies of a scholar is the individual's dissertation. The continuation of research beyond this initial demonstration establishes each individual's position as a scholar within the discipline and his chosen specialty. The demonstration of scholarly competence is not a matter of a single product, but is a continuing event.

"What is your specialty?" is a commonplace question among mathematicians. The question is intended to elicit the particular research interest or specialty of the mathematician. Each research mathematician is expected to have a specialty. Such a question is seldom, if ever, asked of a mathematics educator. Apparently there is little or no research expectation for a mathematics educator.

If a professor is expected to be able to assist others in becoming researchers, then being a researcher in that discipline himself is an obvious prerequisite. Observation suggests that each successful producer of researchers is a researcher himself. This observation appears to hold for every discipline. Competent researchers in multiple disciplines are rare. A competent mathematics researcher should not be expected to be a competent mathematics education researcher. The day of the polymath is past.

The substance of the training that would more likely enable a mathematics educator to make research contributions to his discipline is a program component either overlooked or naively represented by most who describe degrees in mathematics education. One strategy is to omit the research component. This ignores the mathematics educator's necessary role as a scholar. Some programs attempt to accomplish the research training through courses in applied statistics, educational measurement, experimental design, and summaries of various learning theories. This is a naive solution. The consequences of such training are obvious when one examines mathematics education dissertations. Too often the dissertation concerns the investigation of a trivial problem cloaked in an elegant statistical design. This is not to argue that courses in statistics and measurement are unnecessary. Certain of the competencies acquired in these courses are needed, but they are not sufficient for a mathematics education researcher.

Some, especially in these times of increasing competition for available univer-

sity positions, are being as ridiculous as to suggest that a newly degreed research mathematician is also competent as a mathematics educator. Such an argument is built upon the false assumption that mathematics training is the necessary and sufficient condition for competence in mathematics education. One of the theses advanced by this paper is that at least two conditions are required for mathematics education competence: (1) training that enables one to contribute to the research knowledge of mathematics instruction and learning and (2) a knowledge of existing mathematics (say in line with the CUPM recommendations for college teachers). One cannot convert his scholarly-research field by merely changing his title any more than he can change his level of mathematics knowledge by changing his title.

Vicarious encounters with research are not viable substitutes for direct involvement. A prospective research mathematician is guided through a graduate program which applies the principle that if one is to become a researcher, then his research involvement must be at the participating level rather than the spectator level. Clearly, doctoral programs in other disciplines such as the biological sciences and the physical sciences adhere to this principle. In many subject areas proseminars are organized to systematically provide early research experiences for the graduate student. As the graduate student matures, the amount of time allocated to independent research opportunities increases. By the time an individual is working on his dissertation, he has had substantial research experience. Similar active, participating research experiences must become part of a mathematics education doctoral program, if the products of such programs are expected to do research in their chosen field.

It should *not* be expected that each research must produce results that resolve a major problem in mathematics education, but the thesis advanced here is that each investigation be a contribution to the knowledge of the discipline. The majority of research published in any discipline fails to resolve a major problem. It is simply not the characteristic of research to make giant strides with each investigation. Rather, the pace at which research accumulates knowledge is slow but deliberate, with only an occasional breakthrough. This is the nature of research and should not be surprising to anyone, especially to researchers. Research in mathematics education cannot be one of anticipating that most publications will resolve major problems in mathematics learning or instruction. The extent of each contribution and the quality of research will vary in mathematics education as it does in mathematics and every other discipline. The doctoral program in mathematics education should seek to maximize the likelihood that the doctoral student acquires all of the competencies prerequisite to his becoming a researcher.

It should be noted that a mathematics education researcher is distinguished here from a research mathematician. This condition, however, does not exclude the mathematics educator from becoming a researcher. It only restricts any expectation of his becoming a research mathematician. The mathematics educator must be a scholar and an active researcher, if he is to be an effective member of a university faculty.

The instructional activities of a mathematics educator involve preparing mathematics teachers for pre-university positions, assisting current teachers in acquiring additional preparation, improving instructional practices employed by mathematics teachers including those at the university level, assisting in the construction and selection of instructional programs in mathematics, and guiding the training of new researchers in mathematics education. His research interests, therefore, involve the investigation of conditions that affect instruction in mathematics; the development of methodologies of research in mathematics education; and the description of psychological, logical, and chemical mechanisms involved in the learning of mathematics.

2. **Training for Research—How?** Methods of producing researchers are idiosyncratic. Each research professor's methods evolve from his research, experience, successes, and failures. Although each set of procedures is highly individualistic, comment and criticism from one's colleagues is usually beneficial.

At the University of Maryland there is a sequence of seminars that are designed to help the graduate student develop research competencies. The seminars are constructed to involve the student in active research; not merely to expose him to research through vicarious experiences. The doctoral dissertation in mathematics education is not the student's first exposure to research, but is usually a continuation and expansion upon one of the themes to which he was exposed during his graduate work. Below is listed the set of goals for the introductory seminar in the sequence. The purpose of this initial seminar is to introduce the student to a behavioral approach to research on instruction as modeled after the works of Gagné [1], Mager [5], Tyler [7], and Walbesser [8, 9].

The behaviors to be acquired during the seminar are named in Table One. The seminar is hierarchically organized. For example, acquisition of the behavior named in cell 11 (see Table One) is enhanced if the learner has first acquired the behaviors named in cells 9 and 10. Similarly, acquisition of the behavior named in cell 9 is enhanced if the learner has first acquired the behaviors named in cells 8 and 7. Validity estimates on each of these dependencies have been established at the 0.90 level before the hierarchy was introduced as the instructional outline for the seminar. The complete learning hierarchy is described by Chart One.

The learning hierarchy model for investigating instructional variables was selected for several reasons. First, the model is highly generalizable for instructional research at all levels of mathematics teaching. Second, there is an established body of existing knowledge upon which to build. Third, the model is elegant in that it enables the researcher to investigate the effects of one instructional variable on another with comparative ease and yet the assumptions of the model are simple to master.

TABLE ONE

The most complex behavior is described by the statement numbered 38.

1. Identify descriptions of observable performance.
2. Construct definitions for classes of performance which can be exhibited in a variety of settings (e.g., picking up, building).
3. Construct tasks illustrative of a class of performances.

4. Identify each of the six components in a behavioral objective, given Walbesser's definition [9] of a behavioral objective and a statement of a behavioral objective.

5. Describe the meaning of each of the six components in Walbesser's definition of a behavioral objective.

6. Distinguish between the statement of a behavioral objective and a non-behavioral objective.

7. Demonstrate the procedures used in the construction of a set of action verbs and construct definitions for each of the verbs selected, given a set of instructional activities for which behavioral objectives are to be constructed.

8. Demonstrate the use of a set of action verbs in completion of statements describing various human performances.

9. Identify the subordinate behaviors, given a hypothesis of learning dependency.

10. Identify the terminal behavior, given a hypothesis of learning dependency.

11. Identify each hypothesis of learning dependency in a given learning hierarchy.

12. Identify the hypothesis of learning dependency, given a particular terminal behavior and a learning hierarchy.

13. Name the hypothesis of learning dependency, given a particular terminal behavior and a learning hierarchy.

14. Identify behavioral objectives in performance agreement with a given task.

15. Identify tasks in performance agreement with a given behavioral objective.

16. Construct a definition of an assessment task for a given behavioral objective.

17. Identify instructional activities in performance agreement with a behavioral objective.

18. Identify behavioral objectives in performance agreement with an instructional activity.

19. Construct the statement of a behavioral objective for an instructional activity.

20. Construct individualized instructional activities in performance agreement with any of the action verbs.

21. Construct the statement of a behavioral objective and an assessment task given the statement of a non-behavioral objective.

22. Cònstruct behavioral objectives and an assessment task for a behavioral objective.

23. Construct an instructional activity and assessment task for a behavioral objective.

24. Construct a hypothesis of learning dependency, given a description of the terminal behavior.

25. Demonstrate how to find the completeness ratio, given performance data and the hypothesis of learning dependency.

26. Demonstrate how to find the adequacy ratio and the inverse adequacy ratio, given performance data and the hypothesis of learning dependency.

27. Demonstrate how to find the consistency ratio and the inverse consistency ratio, given performance data and the hypothesis of learning dependency.

28. Identify the hypothesis being tested with the adequacy ratio and the hypothesis being tested by the inverse adequacy ratio, given the terminal behavior and the subordinate behaviors.

29. Identify the hypothesis being tested with the consistency ratio and the inverse consistency ratio, given the terminal behavior and the subordinate behaviors.

30. Identify data which indicate reversals on a learning hierarchy, missing subordinate behaviors, inadequate item construction, ineffective instructional activities.

31. Construct revisions in a learning hierarchy based upon decisions arrived at from performance data on the various hypothesis of learning dependency.

32. Demonstrate the collection of data for testing a learning hypothesis in a learning hierarchy.

33. Demonstrate how to validate one hypothesis of learning dependency, given a hypothesis with one subordinate behavior (repeat for two or three subordinate behaviors).

34. Construct a learning hierarchy, given the statement of the terminal behavior tor the hierarchy.

35. Construct the assessment tasks to accompany each cell of the learning hierarchy constructed by the student.

36. Construct instructional activities to accompany each cell of a learning hierarchy constructed by the student.

37. Demonstrate the reporting of the validation of the student's hierarchy in the form of a research report acceptable for publication by a research journal and with sufficient detail to make possible the replication of the investigation.

38. Construct a revision of the learning hierarchy constructed by the student based upon the performance data collected.

Following is listed a brief description of other seminars in the sequence:

SEMINAR TWO: PROBLEM DEFINITION. Purpose: Given data in terms of descriptive observations, tables, or graphs; seminar participants identify sources of variance and construct possible explanations for the observed variance. Explanations are stated in terms of a manipulated variable's possible effect upon a responding variable. Techniques for manipulating variables, obtaining measures for various responding variables, and holding variables constant are demonstrated.

SEMINAR THREE: ANALYSIS AND REPLICATION OF EXISTING RESEARCH. Purpose:

(1) to replicate an experiment reported in the literature and compare findings;

(2) to construct alternative explanations for a given set of findings;

(3) to distinguish between findings and conclusions;

(4) to construct and execute an original research with other members of a team (two or three individuals), given a research hypothesis, and to report the findings by collectively preparing an article for journal publication.

SEMINAR FOUR: INTRODUCTION TO COMPUTER ASSISTED INSTRUCTION, COMPUTER MANAGED INSTRUCTION, AND PROBLEM SOLVING STRATEGIES. Purpose: This seminar provides the student with the opportunity to examine human-computer interactive systems on four levels: data processing, drill and practice, simulation, and problem solving, as defined by Robert M. Gagné [1]. Two group research projects and two individual research projects are incorporated in this seminar.

One purpose in describing such a sequence of seminars is to outline one program for training mathematics education researchers. Another purpose is to invite constructive criticism and suggestions for improvement from the mathematics research community as well as from other mathematics education researchers. The long and successful experience of mathematicians at producing researchers should yield much valuable advice for mathematics education researchers. Finally, it is hoped that this note will stimulate other mathematics educators to share their ideas concerning the research expectations associated with the bi-disciplinary degree of mathematics education.

References

1. R. M. Gagné, The Conditions of Learning, Holt, Rinehart, and Winston, New York, 2nd ed. 1970.

2. I. N. Herstein, On the Ph.D. in mathematics, this MONTHLY, 76 (1969) 818.

3. D. A. Johnson, The doctorate in mathematics education, School Science and Mathematics, 68 (1968) 811.

4. J. M. Laible, A new graduate degree in mathematics, this MONTHLY, 76 (1969) 687.

5. R. F. Mager, Preparing Instructional Objectives, Fearon Publishers, Palo Alto, Calif. 1962.

6. J. M. Scandura, Letter to the editor, The Mathematics Teacher, 60 (1967) 714.

7. R. W. Tyler, Basic Principles of Curriculum and Instruction, University of Chicago Press, 1950.

8. H. H. Walbesser, An Evaluation Model and its Application, American Association for the Advancement of Science, Misc. Pub. 68–4 (1968).

9. ———, Constructing Behavioral Objectives, The Maryland Book Exchange, College Park, Md., 1970.

PROBLEMS AND SOLUTIONS

All problems (both elementary and advanced) proposed for inclusion in this Department should be sent to E. P. Starke, 1000 Kensington Ave., Plainfield, N. J. 07060. Proposers of problems are urged to enclose any solutions or information that will assist the editors. Ordinarily, problems in well-known textbooks and results in generally accessible sources are not appropriate for this Department. No solutions (except those accompanying proposals) should be sent to Professor Starke.

ELEMENTARY PROBLEMS

Solutions of Elementary Problems should be sent to Problems Group, Mathematics Department, University of Maine, Orono, ME 04473. To facilitate their consideration, solutions of Elementary Problems in this issue should be typed (with double spacing) and should be mailed before September 30, 1971. Contributors (in the United States) who desire acknowledgement of receipt of their solutions are asked to enclose self-addressed stamped postcards.

An asterisk (*) means neither the proposer nor the editors supplied a solution.

E 2301. *Proposed by David Singmaster, Bedford College, University of London*

Let G be a group, written additively. Define: (X, d) is a G-directed distance space if d is a function from $X \times X$ to G such that: (1) $d(x, y)=0$ if and only if $x=y$; (2) $d(x, y)=-d(y, x)$; (3) $d(x, z)=d(x, y)+d(y, z)$. Describe all G-directed distance spaces. (X is nonempty.)

E 2302. *Proposed by Erwin Just, Bronx Community College*

Each entry a_{ij} of an nth order square matrix A is the integer $i+j \pmod{n}$. A set of n elements is selected from A so that no two elements appear in the same row, or in the same column. Prove that these n elements can be distinct if and only if n is odd.

E 2303. *Proposed by Charles Lindner, Auburn University*

Let G be a finite group of odd order. Then the set of products of all elements of G, taken in any order, is in the commutator subgroup. [As a corollary we have the well-known result that G abelian implies $\sum_{x\in G} x = \text{identity}$. Query: Does the set of all such products exhaust the commutator subgroup?—*Ed.*]

E 2304. *Proposed by J. C. Owings, Jr., University of Maryland*

Let G be a finite group, H a subgroup of G. Show that, given any left coset L of H, there exists an integer k such that, for any right coset R of H, $L\cap R$ is either empty or has cardinality k.

E 2305*. *Proposed by M. D. Hendy, University of New England, Australia*

In the system of reduced residues modulo p, where p is a prime, for each $e \mid p-1$, there are $\phi(e)$ elements of order e. Prove that the sum of these $\phi(e)$ elements $\equiv \mu(e) \pmod{p}$, where $\phi(e)$ is the Euler totient function and $\mu(e)$ is the Möbius function.

E 2306. *Proposed by Anon, Erehwon-upon-Wabash*

Let A be an $n\times n$ matrix, u a $1\times n$ row vector, v an $n\times 1$ column vector, and

$$B = \begin{bmatrix} A & -Av \\ -uA & uAv \end{bmatrix}.$$

(a) Prove that 0 is a characteristic root of B.
(b) Suppose $\det A = 0$. Show that t^2 divides the characteristic polynomial $\det(tI-B)$ of B.
(c) Discuss the converse of (b).

SOLUTIONS OF ELEMENTARY PROBLEMS

Point in a Cube

E 2103 [1968, 671; 1969, 694]. *Proposed by Simeon Reich, Haifa, Israel*

Find the seven smallest numbers a_k ($k=1, 2, \cdots, 7$) with the following property: If a point P is inside a unit cube $A_1A_2 \cdots A_8$, at most k of the eight distances PA_j ($j=1, 2, \cdots, 8$) are greater than a_k. (Thus at most one of these distances will be greater than a_1, at most two will be greater than a_2, etc.)

II. *Comment by the proposer.* It seems that the published solution does not contain the answer to the proposed problem, where P is *any* point, but answers instead the following question: Find the seven smallest numbers a_k with the property that *there exists* a point P inside a unit cube $A_1A_2 \cdots A_8$ for which

the statement, *At most k of the eight distances PA_j $(j=1, 2, \cdots, 8)$ are greater than a_k*, is true.

III. *Solution by H. V. Monks, Northeastern State College.* Let the distances PA_j be denoted by d_k $(k=1, 2, \cdots, 8)$ where we reorder them if necessary so that $d_1 \geqq d_2 \geqq \cdots \geqq d_8$. It can be established that

$$\begin{array}{ll} \sqrt{3}/2 \leqq d_1 \leqq \sqrt{3}, & \sqrt{2}/2 \leqq d_6 \leqq d_5 \leqq \sqrt{5}/2, \\ \sqrt{3}/2 \leqq d_2 \leqq 3/2, & 1/2 \leqq d_7 \leqq 1, \\ \sqrt{3}/2 \leqq d_4 \leqq d_3 \leqq \sqrt{2}, & 0 \leqq d_8 \leqq \sqrt{3}/2. \end{array}$$

It follows then that the solution is found by taking

$$a_1 = 3/2,\ a_2 = a_3 = \sqrt{2},\ a_4 = a_5 = \sqrt{5}/2,\ a_6 = 1,\ a_7 = \sqrt{3}/2.$$

A Functional Equation

E 2176 [1969, 554; 1970, 310; 1970, 767]. *Proposed by R. S. Luthar, University of Wisconsin at Waukesha*

Find all continuous real functions f such that

$$f\left(\frac{x+y}{x-y}\right) = \frac{f(x)+f(y)}{f(x)-f(y)}\,. \tag{1}$$

II. *Editorial Note.* In his solution, Simeon Reich (Haifa, Israel) points out that the only real function, continuous or not, which satisfies the given equation is the identity function. The gist of his argument follows. Since $f(zx)=f(z)f(x)$, it follows that $f(y)=(f(\sqrt{y}))^2>0$ for all $y>0$. Now

$$f(x) - f(y) = \frac{f(x)+f(y)}{f\left(\dfrac{x+y}{x-y}\right)} > 0$$

for all x, y with $x>y\geqq 0$. Hence f is order-preserving on the positive reals. Since f fixes the positive rationals, it must fix the positive reals. For, suppose there exists $x>0$ with $f(x)>x$; choose a rational r with $x<r<f(x)$. Then $f(x)<f(r)=r<f(x)$. Similarly it cannot be the case that $f(x)<x$. Finally, since $f(-x)=-f(x)$, it follows that f is the identity function.

A Comparison of Integrals

E 2216 [1970, 192; 1114]. *Proposed by M. S. Klamkin, Ford Scientific Laboratory*

Which of the two integrals

$$\int_0^1 x^x\,dx, \qquad \int_0^1\int_0^1 (xy)^{xy}\,dx\,dy$$

is larger?

II. *Comment by C. D. Olds, San Jose State College.* For readers who wonder how the computer value reported by Klein might be obtained, the following manipulations (easily justified) may be of interest.

$$I = \int_0^1 x^x dx = \int_0^1 e^{x \ln x}\, dx = \int_0^1 \sum_{n=0}^{\infty} \frac{(x \ln x)^n}{n!}\, dx$$

$$= \sum_{n=0}^{\infty} \frac{1}{n!} \int_0^1 (x \ln x)^n dx = \sum_{n=0}^{\infty} \frac{(-1)^n}{(n+1)^{n+1} n!} \int_0^{\infty} e^{-t} t^n dt$$

$$= \sum_{n=0}^{\infty} \frac{(-1)^n}{(n+1)^{n+1} n!} \Gamma(n+1) = \sum_{n=0}^{\infty} \frac{(-1)^n}{(n+1)^{n+1}} = 0.78343051 \cdots.$$

The series is particularly attractive because of its rapid convergence.

A Condition which Implies Monotonicity

E 2246 [1970, 653]. *Proposed by Andrew Rochman, Saint Louis University*

Let f be a nonconstant, real valued, continuous function such that, for all $x, y \in R$, $f(x+y) = \phi(f(x), y)$. Prove that f is monotonic.

I. *Solution by G. A. Heuer, Concordia College.* Indeed, f is strictly monotone or constant. For, suppose $f(a) = f(b)$ for some a, b with $b-a=h>0$. Then $f(a+t) = \phi(f(a), t) = \phi(f(b), t) = f(b+t)$ for all t; i.e., f is periodic with period h. Since $f(a) = f(b)$, there is a point c in (a, b) such that $f(c)$ is an extreme value on $[a, b]$. If $\epsilon > 0$ is chosen so that $c-\epsilon$ and $c+\epsilon$ are in $[a, b]$, then $f(x) - f(x+\epsilon)$ has opposite signs at $x=c$ and $x=c-\epsilon$, so is zero at some point a_0 between; thus $f(a_0) = f(a_0+\epsilon)$. But this implies f is periodic with period ϵ. Hence f is constant.

II. *Solution by K. S. Sarkaria, State University of New York at Stony Brook.* Suppose f is not monotonic; then we can find three points $x_1 < x_2 < x_3$ such that $f(x_1) > f(x_2) < f(x_3)$ or $f(x_1) < f(x_2) > f(x_3)$. We can assume the first case since the proof for the second case is similar.

Now choose a real number A so that $f(x_2) < A < \min\{f(x_1), f(x_3)\}$, and define a and b by

$$a = \text{lub}\{x \mid f(x) = A, x < x_2\}, \qquad b = \text{glb}\{x \mid f(x) = A, x > x_2\}.$$

Obviously we have $x_1 \leqq a < x_2 < b \leqq x_3$ and the function $f(x)$ is bounded above by A in $[a, b]$. Further, since $f(a) = f(b) = A$, it follows that $f(x+b-a) = \phi(f(b), x-a) = \phi(f(a), x-a) = f(x)$, for all x. Hence $f(x)$ is periodic with period $b-a$ and so is bounded above by A for all values of x. This is not possible.

Also solved by J. Aczel, B. H. Aupetit, Donald Batman, R. L. Enison, Michael Greening, Eric Langford, Harry Lass, William Margolis, Robert Patenaude, M. A. Radke, Wayne Roberts, Sid Spital & Peter Fowler, D. P. Stanford, Walter Stromquist, and B. L. D. Thorpe.

Professor Aczel notes that the problem is not new and that its solution can be found in Aczel, Kalmar and Mikusinski, *Sur l'équation de translation*, Studia Math. 12 (1951), 112–116. The prob-

lem is also solved in his book, *Lectures on Functional Equations and their Applications*, Academic Press, New York, 1966, p. 19. He has generalized the problem in a later paper, *On strict monotonicity of continuous solutions of certain types of functional equations*, Canad. Math. Bull., 9 (1966), 229–232.

Representation of Integers in Terms of a Given Set

E 2247 [1970, 765]. *Proposed by N. S. Mendelsohn, University of Manitoba*

Let $a_1, \cdots, a_n$ be a set of relatively prime positive integers. Let $F(a_1, \cdots, a_n)$ represent the largest integer which cannot be represented in the form $c_1a_1+c_2a_2+\cdots+c_na_n$, where $c_1, \cdots, c_n$ are nonnegative integers. Prove the following:

(1) If $(a, b)=1$, $a>0$, $b>0$, and c is a positive integer such that c is nonrepresentable in the form $Aa+Bb$, with A, B nonnegative integers, then $F(a, b, c)<F(a, b)$.

(2) If $(a, b)=1$ and $2<a<b$ and t is an integer such that $ta<b<(t+1)a$, then $F(a, b, ab-(t+1)a-b)=ab-a-2b$.

Solution by the proposer. Let m and n be positive integers such that $(m, n)=1$.

LEMMA. *If A and B are integers such that $A+B=mn-m-n$ then exactly one of A and B is representable in terms of m and n.*

Proof: If both A and B were representable then $A=rm+sn$, $B=tm+un$, so that $mn-m-n=(r+t)m+(s+u)n$. Hence $mn=(r+t+1)m+(s+u+1)n$. This implies $(s+u+1)\equiv 0 \bmod m$ and $(r+t+1)\equiv 0 \bmod n$. But since $s+u+1>0$ and $r+t+1>0$, we have $s+u+1\geqq m$ and $r+t+1\geqq n$ so that $mn=(r+t+1)m+(s+u+1)n\geqq 2mn$, a contradiction. Now suppose A is not representable. Since $(m, n)=1$, $A=Rm-Sn$ where $S>0$ and $0\leqq R\leqq n-1$. Hence $B=mn-m-n-A=mn-m-n-Rm+Sn=(n-R-1)m+(S-1)n$ so that B is representable.

COROLLARY. *Since 0 is representable but no negative integer is representable, $F(m, n)=mn-m-n$.*

Proof of (1): Let $(a, b)=1$. If c is not representable in terms of a and b, then by the lemma, $ab-a-b-c$ is representable, i.e., $ab-a-b-c=ra+sb$ where r and s are nonnegative integers. Therefore $ab-a-b=ra+sb+c$. It follows that $F(a, b, c)<F(a, b)$.

Proof of (2): By repeated applications of the lemma it is seen that the following integers are not representable in terms of a and b: $ab-a-b, ab-2a-b, \cdots, ab-ta-b, ab-a-2b$. Now $ab-ra-b=\{ab-(t+1)a-b\}+(t+1-r)a+0b$ is representable in terms of a, b, $\{ab-(t+1)a-b\}$ for $0\leqq r\leqq t+1$. However, if $ab-a-2b$ were also so representable, then $ab-a-2b=Aa+Bb+C\{ab-(t+1)a-b\}$ with A and B nonnegative. Now $C\neq 0$ because $ab-a-2b$ is not representable in terms of a and b alone. On the other hand $C=0$ is implied by $ta<b$ whence $ab-a-2b<ab-(t+1)a-b$. From this contradiction the desired conclusion follows.

Also solved by Paul Fan, N. Felsinger, M. G. Greening (Australia), and Simeon Reich (Israel).

A Triangle Triple Inequality

E 2248 [1970, 765]. *Proposed by A. W. Walker, Toronto, Canada*

If a, b, c, r, R are the side lengths, inradius and circumradius of any plane triangle, then

$$\frac{1}{2rR} \leqq \frac{1}{3}\left(\frac{1}{a}+\frac{1}{b}+\frac{1}{c}\right)^2 \leqq \frac{1}{a^2}+\frac{1}{b^2}+\frac{1}{c^2} \leqq \frac{1}{4r^2}\cdot$$

Solution by Henrik Meyer, Birkerød, Denmark. The inequality

$$(1/a - 1/b)^2 + (1/b - 1/c)^2 + (1/c - 1/a)^2 \geqq 0$$

is equivalent to the following two inequalities

$$1/bc + 1/ca + 1/ab \leqq \tfrac{1}{3}(1/a + 1/b + 1/c)^2,$$
$$\tfrac{1}{3}(1/a + 1/b + 1/c)^2 \leqq 1/a^2 + 1/b^2 + 1/c^2.$$

From the theorem about the arithmetic and geometric means we get

$$(s-a)(s-b) \leqq \left(\frac{s-a+s-b}{2}\right)^2 = \frac{c^2}{4}$$

and similar formulas. Now

$$\begin{aligned}1/2rR &= 2s/abc = (a+b+c)/abc = 1/bc + 1/ca + 1/ab \\ &\leqq \tfrac{1}{3}(1/a + 1/b + 1/c)^2 \leqq 1/a^2 + 1/b^2 + 1/c^2 \\ &\leqq \tfrac{1}{4}\{1/(s-b)(s-c) + 1/(s-c)(s-a) + 1/(s-a)(s-b)\} \\ &= \frac{s-a+s-b+s-c}{4(s-a)(s-b)(s-c)} = \frac{s}{4(s-a)(s-b)(s-c)} = 1/4r^2.\end{aligned}$$

Also solved by Alfred Aeppli, W. J. Blundon, L. Carlitz, C. S. Karuppan Chetty (India), G. V. Ferrer (Mexico), Ralph Garfield, Michael Goldberg, M. G. Greening (Australia), Robert Heller, Toshi Iida (Japan), Peter Kornya, Norman Miller, G. R. Padmanabhan & T. V. Lakshminarasimhan & K. Rangarajan (India), A. J. Patsche, Simeon Reich (Israel), S. Shahabi (Iran), David Spear, John Steinig, Jim Tattersall, Al Wiener, and the proposer.

ADVANCED PROBLEMS

All solutions of Advanced Problems should be sent to J. Barlaz, Rutgers—The State University, New Brunswick, N. J. 08903. Solutions of Advanced Problems in this issue should be typed (with double spacing) on separate, signed sheets and should be mailed before September 30, 1971. Contributors (in the United States) who desire acknowledgement of receipt of their solutions are asked to enclose self-addressed, stamped postcards.

An asterisk (*) means neither the proposer nor the editors supplied a solution.

5802.* *Proposed by J. P. Jones, University of Calgary*

The Cantor set X has the property that for every positive real number d, X contains points x_0, x_1 such that $d = x_1 - x_0$. More generally, does there exist a set X

of measure zero such that for every finite sequence $d_1, d_2, \cdots, d_n$ of positive real numbers, X contains points $x_0, x_1, \cdots, x_n$ such that $d_i = x_i - x_{i-1}$ for $i=1, 2, \cdots, n$?

5803. *Proposed by G. Sabbagh, Yale University*

Let A be a ring with 1 and without zero divisors. It is obvious that, if A is commutative, then the torsion elements of each left A-module E constitute a submodule of E. What other rings have this property?

5804. *Proposed by D. A. Herrero, University of Chicago*

Let $L(H)$ denote the space of all bounded linear operators on the Hilbert space H. Prove that the closed convex hull of the set of all unitary operators is dense in the closed unit hull of $L(H)$.

5805.* *Proposed by John Stout, New York, N. Y.*

For any $x \in (0, 1]$, there is exactly one binary expression $x = 0.\, x_1x_2x_3 \cdots = \sum_{i=1}^{\infty} x_i\, 2^{-i}$ with an infinite number of ones. Define a function f: $(0, 1] \to R^+ \cup \{+\infty\}$ by $f(x) = \sum_{i=1}^{\infty} x_i/i$. Let C be the inverse image of R^+ and D the inverse image of $\{+\infty\}$.

Are C and D Lebesgue measurable? If so, what are their measures?

5806. *Proposed by Dennis Allen, Jr., Michigan Technological University, Houghton*

Let G denote the ring of Gaussian integers and $G[[x]]$ the ring of formal power series over G. Let $a_1, a_2, \cdots, a_n$ be Gaussian integers, each with positive real part, and let $e_1, \cdots, e_n$ be whole numbers. Suppose

$$\prod_{k=1}^{n}\left[\sum_{j=0}^{\infty}(a_kx)^j\right]^{e_k} = \sum_{j=0}^{\infty} b_jx^j.$$

Does it follow that $b_j \neq 0$ for $0 \leqq j \leqq \min\{e_1, \cdots, e_n\}$?

5807. *Proposed by E. G. Kundert, University of Massachusetts*

Put

$$\alpha(s; h, k) = \binom{s}{k}\binom{k}{s-h} = \frac{s!}{(s-h)!\,(s-k)!\,(h+k-s)!}.$$

Let p be a fixed prime number and i, j fixed natural numbers. Prove

$$\sum \alpha(s_1; i, i)\alpha(s_2; i, s_1) \cdots \alpha(s_{p-2}; i, s_{p-3})\alpha(i+j; i, s_{p-2}) \equiv 0 \pmod{p},$$

where the summation is over all $s_1, s_2, \cdots, s_{p-2}$ such that

$$i \leqq s_1 \leqq s_2 \leqq \cdots \leqq s_{p-2} \leqq i+j.$$

SOLUTIONS OF ADVANCED PROBLEMS

Non-linear Solutions of $g'(cx)=c[g(x+1)-g(x)]$

5718 [1970, 198]. *Proposed by M. T. Boswell and G. P. Patil, Pennsylvania State University*

Find all functions $g(x)$ which are complex-valued functions of a real variable satisfying

$$\frac{dg(cx)}{dx} = c[g(x+1) - g(x)]$$

for $x>0$ and fixed $c>0$. The linear functions satisfy this equation.

Partial Solution by R. D. Meredith, Stanford University. By transforming the variable it is easy to see that $g(x)$ has derivatives of all orders. Using techniques developed in the note, R. A. Horn and R. D. Meredith, *On a functional equation arising in probability*, this MONTHLY 76 (1969), 802–804, we are able to prove:

THEOREM I. *If g satisfies the given equation and*

(i) *g has a real analytic extension to the unbounded interval* $(-1-\epsilon, \infty)$ *for some* $\epsilon>0$,

(ii) *the Maclaurin series for g has radius of convergence* >1,

(iii) $\sum_{n=0}^{\infty} g^{(n)}(0)/\lambda_0^n<\infty$ *for some* $\lambda_0>0$ *satisfying* $\lambda_0<\mu_0$ *if* $c<1$, *where* $e^{\mu_0}-\mu_0=2$, *and* $\lambda_0<\mu_1$ *if* $c=1$, *where* $e^{\mu_1}-\frac{1}{2}\mu_1^2-\mu_1=3/2$,

then g is linear.

THEOREM 2. *If g satisfies the given equation with* $c>1$, *and if for some* $\alpha>0$ *we have* $\lim_{x\to\infty} x^\alpha g(x)=0$, *then g is linear.*

An example of a nonlinear function satisfying the given differential equation when $c=1$ is $g(x)=e^{ax}$ where a is a nonreal root of $e^a=a+1$.

Density of Pairs with Same Prime Factors

5735 [1970, 532]. *Proposed by Paul Erdös and T. Motzkin, University College of Swansea, Wales*

Denote by $F(n)$ the number of pairs $1\leqq a\leqq b\leqq n$ for which a and b have the same prime factors. Prove that $\lim_{n\to\infty} F(n)/n$ exists and is finite.

Solution by R. T. Bumby, Rutgers—The State University. The limit is $\zeta(2)\zeta(3)/\zeta(6)$.

If $n=r_1^{a_1}r_2^{a_2}\cdots r_k^{a_k}$ with each r_i prime and $a_i\geqq 1$, the difference

$$f(n) = F(n) - F(n-1) = a_1a_2\cdots a_k$$

is clearly a multiplicative number theoretic function. We first obtain another counting expression for $f(n)$. Let S be the set of integers m (including 1) each of whose prime factors has exponent more than one. Let $c(m)$ be the characteristic function of S. Then $c(m)$ is multiplicative.

Let $g(n)=\sum_{m|n} c(m)$; $g(n)$ is multiplicative and $g(p^\alpha)=\alpha=f(p^\alpha)$, p prime; and so $f(n)=g(n)$ for all n. Therefore

$$F(n) = \sum_{k=1}^{n} g(k) = \sum_{k=1}^{n} \sum_{m|k} c(m) = \sum_{m\in S, m\leq n} \left[\frac{n}{m}\right].$$

The set S has density zero. (See Niven and Zuckerman, *Introduction to the Theory of Numbers*, Theorem 11.7.) Moreover, $\sum 1/m$ converges and from the identity $\zeta(n) = \prod_p (1-p^{-n})^{-1}$ we get

$$\sum_{m\in S} \frac{1}{m} = \prod_p \left(1 + \frac{1}{p^2} + \frac{1}{p^3} + \cdots\right) = \prod_p \left(1 + \frac{1}{p(p-1)}\right) = \frac{\zeta(2)\zeta(3)}{\zeta(6)}\cdot$$

With $K = \zeta(2)\zeta(3)/\zeta(6)$ we have

$$\left|\frac{Kn - F(n)}{n}\right| \leqq \frac{1}{n} \sum_{m\in S} \left(\frac{n}{m} - \left[\frac{n}{m}\right]\right) \leqq \frac{1}{n} \sum_{m\leq n} c(m) + \sum_{m\in S, m>n} \frac{1}{m}\cdot$$

The first sum on the right tends to zero because S has zero density; the second tends to zero because of the convergence of $\sum c(m)/m$.

The form of proof used above is discussed in the paper, F. T. Atkinson and L. Cherwell, *The mean values of arithmetical functions*, Quarterly Journal of Mathematics 20 (1949), p. 65, ff. and is attributed to A. Axer in a paper in the *Wiener Sitzungsberichte* 120 (1911).

Locating Zeros of Polynomials

5737 [1970, 532]. *Proposed by Simeon Reich, Israel Institute of Technology*

Let $p(z) = a_0 + a_1 z + \cdots + a_{n-1}z^{n-1} + a_n z^n$, $a_n \neq 0$, be a polynomial with complex coefficients. Denote

$$\left\{ \max_{0\leqq k\leqq n-1} |a_k/a_n| \right\}^{1/n}$$

by Q. It is known (Morris Marden, *The Geometry of the Zeros*, AMS, 1949, p. 96) that all zeros of $p(z)$ lie in the circle $|z| \leqq 1 + Q^n$. Suppose that $a_{n-1} = 0$, and that $Q > 1$. Prove that the zeros of $p(z)$ lie also in the circles

$$|z| \leqq Q + Q^2 + \cdots + Q^{n-1}, \quad \text{and} \tag{1}$$

$$|z| \leqq \max_{0\leqq j,k\leqq n-1, j\neq k} \{(1 + |a_k/a_n|)(1 + |a_j/a_n|)\}^{1/2}. \tag{2}$$

Solution by O. P. Lossers, Technological University, Eindhoven, Netherlands. Without loss of generality we may assume that $a_n = 1$. The conditions $a_{n-1} = 0$, $Q > 1$ imply that $n \geqq 2$.

In the case $n = 2$ we have $p(z) = a_0 + z^2$, $|a_0| = Q^2$, so $|z| = Q$ for each zero z of $p(z)$, and the inequalities (1), (2) are trivially satisfied.

Hence we suppose $n \geqq 3$. We are first going to prove a third inequality for the zeros z of $p(z)$, namely

$$|z| < \tfrac{1}{2} + \sqrt{\tfrac{1}{4} + Q^n}. \tag{3}$$

Let z be a zero of $p(z)$ such that $|z|>1$ (otherwise there is nothing to prove). Since $-z^n=a_0+a_1z+\cdots+a_{n-2}z^{n-2}$ we easily derive that

$$|z|^n = Q^n\frac{|z|^{n-1}-1}{|z|-1},$$

and from this inequality and the fact that $|z|^{n-1}>|z|^{n-1}-1$ it follows that $Q^n/(|z|-1)>|z|$, which result leads to inequality (3). (3) is stronger than (1) because

$$\tfrac{1}{2}+\sqrt{\tfrac{1}{4}+Q^n}<1+Q^{n/2}<Q+Q^2+\cdots+Q^{n-1}.$$

Let $j\in\{0,1,\cdots,n-2\}$ be such that $|a_j|=Q^n$. In considering inequality (2), we find two possibilities:

A. For each $i\in\{0,\cdots,n-2\}$, $i\neq j$, we have $|a_i|\leqq\alpha$, with $\alpha=\sqrt{2}-1$. Let z be a zero of $p(z)$ such that $|z|>1$ (otherwise there is nothing to prove). Then

$$\begin{aligned}|z|^n &\leqq |a_0|+|a_1|\cdot|z|+\cdots+|a_{n-2}|\cdot|z|^{n-2}\\ &\leqq \alpha+\alpha|z|+\cdots+\alpha|z|^{n-3}+Q^n|z|^{n-2},\end{aligned}$$

so

$$\text{(i)}\qquad 1\leqq\frac{\alpha}{|z|^n}+\frac{\alpha}{|z|^{n-1}}+\cdots+\frac{\alpha}{|z|^3}+\frac{Q^n}{|z|^2}\cdot$$

We assert that $|z|<\sqrt{1+Q^n}$, an inequality which implies (2). Since the right-hand side of (i) is a decreasing function of $|z|$, it is sufficient to prove that for $t=\sqrt{1+Q^n}$ we have

$$1>\frac{\alpha}{t^n}+\frac{\alpha}{t^{n-1}}+\cdots+\frac{\alpha}{t^3}+\frac{Q^n}{t^2},$$

or, equivalently,

$$1>\frac{\alpha}{t^{n-2}}+\frac{\alpha}{t^{n-3}}+\cdots+\frac{\alpha}{t}\cdot$$

As $t>\sqrt{2}$, we see that

$$\begin{aligned}\frac{\alpha}{t^{n-2}}+\cdots+\frac{\alpha}{t} &< (\sqrt{2}-1)\left[\frac{\frac{1}{\sqrt{2}}\left\{1-\left(\frac{1}{\sqrt{2}}\right)^{n-2}\right\}}{1-\frac{1}{\sqrt{2}}}\right]\\ &< (\sqrt{2}-1)\frac{1/\sqrt{2}}{1-1/\sqrt{2}}=1,\end{aligned}$$

as desired.

B. There is an $i\in\{0, \cdots, n-2\}$, $i\neq j$ such that $|a_i|>\sqrt{2}-1$. It is sufficient to prove that $|z|<\{(1+Q^n)\sqrt{2}\}^{1/2}$ for each zero z of $p(z)$; the last inequality follows from (3) because for $Q>1$,

$$\tfrac{1}{2}+\sqrt{\tfrac{1}{4}+Q^n}<\{(1+Q^n)\sqrt{2}\}^{1/2}.$$

Also solved by Emeric Deutsch, A. A. Jagers (Netherlands), Q. G. Mohammad (India), Alberto Torchinsky, Robert Vermes, and the proposer.

Note. Both Deutsch and the proposer introduce a companion matrix for which $p(z)$ is the characteristic polynomial, and then apply theorems which may be found in M. Marcus and H. Mink, *A Survey of Matrix Theory*, pp. 52, 146, ff.

Representation of Primes

5738 [1970, 533]. *Proposed by E. M. Reingold, University of Illinois*

Prove that there are infinitely many primes whose representations in base b begin (on the left) with an arbitrary string of digits $a_na_{n-1}\cdots a_0$ (where $0\leq a_i<b$).

I. *Solution by George Barany, Student, Stuyvesant High School, New York City.* The problem is clearly equivalent to proving an infinity of primes of the form

$$p=Ab^n+K,\qquad K<b^n,\qquad A,\,b\text{ fixed.}$$

A consequence of the prime number theorem is that for sufficiently large x and arbitrary $\epsilon>0$, there is a prime between x and $(1+\epsilon)x$ (further, $\lim_{x\to\infty}\pi(1+\epsilon)x-\pi(x)=\infty$). With $\epsilon=1/A$, K can be chosen to make p prime for all sufficiently large n.

II. *Solution by E. W. Trost, Technikum Winterthur, Switzerland.* W. Sierpiński (Acta Arithmetica 5 (1959), 265–266) has proved the following theorem: Let $a_1, a_2, \cdots, a_m$ and $b_1, b_2, \cdots, b_n$ be two strings of decimal digits, where $b_n=1, 3, 7$, or 9. There exists an arbitrarily large prime number whose decimal representation begins with $a_1a_2\cdots a_m\cdots$ and terminates with $\cdots b_1b_2\cdots b_n$. Sierpiński has pointed out that following a suggestion of S. Knapowski one can prove the theorem for an arbitrary base $b>1$ if the condition $(b_n, 10)=1$ is replaced by $(b_n, b)=1$.

Also solved by Dennis Allen, Jr., W. M. Conner, Leon Gerber, David Kelly, M. S. Klamkin, Andrzej Makowski (Poland), A. M. Vaidya & V. S. Joshi (India), and the proposer.

Abelian Groups with Connected Dual Group

5739 [1970, 533]. *Proposed by Richard Johnsonbaugh, Morehouse College*

Find all connected, locally compact abelian groups with connected dual group.

Solution by G. M. Leibowitz, University of Connecticut. Let G be such a group, and let X be its dual group. By the basic structure theorem (see, e.g.,

Theorem 9.14 in Hewitt and Ross, *Abstract Harmonic Analysis*, Vol. 1), G is topologically isomorphic with $R^n \times E$, where n is a nonnegative integer and E is a compact connected group. Hence X is topologically isomorphic with $R^n \times F$, where F is the dual group of E. Since E is compact and X is connected, F is discrete and connected; hence, $F = \{0\}$. So X is isomorphic to R^n and hence by duality, G is isomorphic to R^n.

Also solved by R. W. Chaney, G. A. Edgar, D. A. Hejhal, A. A. Jagers (Netherlands), David Kelly, N. Niederreiter, and the proposer.

A Non-surjective Canonical Map

5740 [1970, 655]. *Proposed by L. R. Etzweiler and W. C. Waterhouse, Cornell University*

Let R be a commutative ring, S a multiplicative subset of R. Let M and N be R-modules. There is a canonical map

$$S^{-1}\operatorname{Hom}_R(M, N) \to \operatorname{Hom}_{S^{-1}R}(S^{-1}M, S^{-1}N).$$

This map is known to be injective if M is finitely generated, and bijective if M is finitely presented (N. Bourbaki, *Algèbre Commutative*, Ch. II, Section 2, Prop. 19). Find an example in which M is finitely generated and the map is not surjective.

Solution by A. A. Jagers, Technische Hogeschool Twente, Enschede, Netherlands. Denote by N the set of all natural numbers and by Z the ring of the integers. Let P be the polynomial ring $Z[X_1, X_2, \cdots]$, and J the ideal of P generated by all products $X_{2k}X_{2k-1}$ $(k \in N)$.

Put $R = P \bmod J$ and $x_i = X_i \bmod J$. Let L_0 be a free R-module with one generator g and let L_1 be the submodule of L_0 generated by all products $x_{2k}g$ $(k \in N)$. Put $M = L_0 \bmod L_1$ and $m = g \bmod L_1$. Let N be equal to R considered as an R-module. Let S be the multiplicative subset of R generated by 1 and all x_{2k-1} $(k \in N)$ and let $\phi: R \to S^{-1}R$ be the canonical homomorphism. Then ϕ is an isomorphism when restricted to the subring generated by S. But $\phi(x_{2k}) = 0$, since $x_{2k-1}x_{2k} = 0$ and $x_{2k-1} \in S$. Hence $S^{-1}R$ is isomorphic to the ring of rational functions in x_{2k-1} $(k \in N)$ with integer coefficients.

One shows in the same way that both $S^{-1}M$ and $S^{-1}N$ are isomorphic to $S^{-1}R$ considered as a $S^{-1}R$-module. Hence

$$\operatorname{Hom}_{S^{-1}R}(S^{-1}M, S^{-1}N) \neq (0).$$

On the other hand if $f \in \operatorname{Hom}_R(M, N)$, then $0 = f(0) = f(x_{2k} \cdot m) = x_{2k} \cdot f(m)$ so that $f(m)$ is a multiple of x_{2k-1}. Since this holds for all k, $f(m)$ must be equal to 0. Hence $f \equiv 0$ because m generates M. Thus $\operatorname{Hom}_R(M, N) = (0)$. It is now evident that R, S, M, and N constitute the desired example.

Also solved by the proposers.

REVIEWS

EDITED BY J. ARTHUR SEEBACH, JR., AND LYNN A. STEEN

with the assistance of the mathematics departments of St. Olaf and Carleton Colleges.

COLLABORATING EDITOR FOR FILMS: SEYMOUR SCHUSTER

Printed materials for review should be sent to: Book Review Editor, American Mathematical Monthly, St. Olaf College, Northfield, MN 55057. Films and correspondence relating to films should be sent to Seymour Schuster, Carleton College, Northfield, MN 55057.
All unsigned material is written by one of the editors. A boldface capital **C** *in the margin indicates that a review is based in part on classroom use. Professors willing to write such a review should inform the editor in order to avoid duplication.*

Applications of Model Theory to Algebra, Analysis, and Probability. Edited by W. A. J. Luxemburg. Holt, Rinehart and Winston, New York, 1969. vii+307 pp. $10.50. (Telegraphic Review, August 1969.)

This volume contains versions, sometimes greatly expanded, of papers read at an International Symposium on its subject in 1967. Their main theme is the applications of nonstandard models. A nonstandard model for, say, the real numbers is a mathematical structure which obeys all the axioms for the real numbers provided that the higher order terms in the language used in these axioms (that is, terms such as set or function) are given interpretations according to rules which, while in agreement with the rules of set theory, do not allow them to apply to all the objects to which we should usually apply them. The existence of such systems was discovered by Skolem in 1934. The modern development of their theory, and the demonstration of its usefulness as a tool in proving results in algebra and analysis, is due to Abraham Robinson.

The articles printed cover a wide range. Robinson discusses applications to algebra and to function theory. In a paper with Zakon he develops a new approach, based on set theory rather than on type theory and logic, to nonstandard models. In a long and very interesting article Luxemburg develops a general theory of monads with numerous applications to topology and functional analysis. In all, there are nineteen papers, ranging from studies in mathematical logic to accounts of applications to topology, analysis and probability theory. The latter are among the less recondite. They include proofs of limiting theorems of probability theory, and an account of how the use of infinitesimal measures enables us to make comparisons between the probabilities of events with zero probability in the ordinary theory.

A minimum requirement for understanding most of the articles is familiarity with the terminology of model theory and mathematical logic. Some of the articles, such as that of Luxemburg, take the subject *ab initio* and should in principle be intelligible to readers with this minimum, but anyone aiming to understand this book would be well advised to acquaint himself with Robinson's standard work beforehand. However, any mathematician could derive some interest from browsing through the book and learning of the applications of this very recent method to branches of the mathematics in which he is interested.

Mathematicians acquainted with the subject will find it invaluable as an account of its development and applications.

J. L. B. COOPER, Chelsea College, London

Exploring University Mathematics 3. Lectures given at Bedford College, London. Edited by N. J. Hardiman. Pergamon Press, Oxford, 1969. x+119 pp. \$3.00 (flexicover), \$4.75 (cloth). (Telegraphic Review, November 1969.)

The seven lectures in this book were delivered at the third of an annual series of conferences for "students about to embark on a degree course of which mathematics is a major part." The lectures are independent of one another. Some presuppose a background in introductory calculus or beginning physics. They range from descriptive summaries to considerations in some depth of limited topics. Among the descriptive lectures, *Mathematics and the Physicist* by M. R. Hoare is worthy of particular mention. It might well be required reading for the budding physicists in a second term calculus course, and it would do the prospective mathematics major no harm to read it also. The two final lectures, *Some Applications of the Taylor Series in Numerical Analysis* by A. Graham, and *Some Irrational Numbers* by H. G. Eggleston, are excellent source materials for lectures by students to a mathematics club or a freshman honors seminar. In each case the exposition is remarkably clear, and yet the student will want and be able to supply greater detail and elaboration by himself and, in so doing, will get a feeling of real accomplishment. There are a few misprints, but an alert student should have no difficulty in discovering them.

J. V. FINCH, Beloit College

Information Utilities. By Richard E. Sprague. Prentice-Hall, Englewood Cliffs, N. J., 1970. 208 pp. \$8.50. (Telegraphic Review, May 1970.)

Although the style is irritating, this book gives a good overview of the extent to which computer utilities are already being used for problem solving, ticketing, banking, etc.; of the rapid rate of growth of these utilities; and of some of the social and economic implications. There is no mathematical content.

C. C. GOTLIEB, University of Toronto

Numerical Approximations to Functions and Data. Based on a Conference Organized by the Institute of Mathematics and Its Applications, Canterbury, England, 1967. Edited by J. G. Hayes. University of London Press, 1970. 177 pp. \$9.60. (Telegraphic Review, October 1970.)

As indicated on the jacket, the aim of this book is to help the user faced with practical problems in two areas of numerical approximation: the approximation of functions and the fitting of curves or surfaces to empirical data. Numerical methods for a variety of problems are given, with perhaps the use of cubic splines predominating, and these methods are illustrated with numerous, but briefly stated, examples.

From the standpoint of the practicing numerical analyst, these methods

would serve to point out possibilities and would direct him to a number of sources which would undoubtedly supply more details concerning the methods and the examples. For instance, M. J. D. Powell, one of the contributors, offers to furnish a Fortran listing of his program for 'Curve fitting by splines in one variable'.

For a text, I feel, it would be most appropriate for a seminar where the participants already have a fairly good background in approximation theory (such as in Cheney—*Introduction to Approximation Theory*), and for some parts of the book, a good knowledge of mathematical statistics. Some of the topics require a rather extensive background in analysis, but again, there are many references furnished for further supplementary reading.

For the researcher, the material presented points to various areas of approximation theory where extensive work is being done.

A. B. FARNELL, Colorado State University

Introduction to Potential Theory. By L. L. Helms. Wiley—Interscience Series in Pure and Applied Mathematics, vol. 22, New York, 1969. ix+282 pp. $14.95. (Telegraphic Review, March 1970.)

Any book with this title would be very welcome and deserve the careful attention of everyone interested in mathematical potential theory. For a long time the only book available in English has been the classical work by O. D. Kellogg, *Foundations of Potential Theory*. This was written in 1929 and has never been revised and so, important though it was, it is by now of little use to anyone wishing for an introduction to this important branch of modern mathematics.

Potential theory today is a very broad subject reaching into such diverse areas as Markov processes, martingales, Riemann surfaces. It makes use of concepts and results arising in these topics as well as in the theory of Hilbert spaces and partial differential equations. The classical theory contained a wealth of valuable concepts such as potentials, energy, harmonic functions, capacity, Green functions, balayage and subharmonic functions. During the second third of this century each of these was given a rigorous mathematical treatment and as a result the various parts of this same field often become difficult to relate. In addition, although the connections of potential theory with partial differential equations and complex variable theory are classical, in recent years the much less obvious but equally important connections with Markov processes have been elucidated. Thus any author setting out to write an introduction to potential theory is faced with a very difficult task; his field is, as indicated, vast, and also his audience is varied. It is clearly impossible to cover all the ramifications of the theory, or even in fact to deal with all of the basic material; a choice has to be made.

The present excellent book has made this choice very well. What has been omitted—except for one area that will be mentioned later—can be considered as second level material. Thus no time is spent on Markov processes, Dirichlet spaces, complex variable theory, partial differential equations or axiomatic

potential theory. Anyone having read this book, however, will have been introduced to all the concepts and basic results in potential theory necessary for these areas. The only introductory material that can be considered as having been omitted is the subject of potential theory with kernels. As a result anyone wishing to find out about recent work in this area, or about the work of the Japanese school of potential theory, will not find the material in the book sufficient although even for these purposes the book is essential reading.

As to the author's intentions and aims, the best thing to do is to quote from his introduction: "In writing *Introduction to Potential Theory,* I have tried to make the material accessible to the student who has just completed basic courses in measure theory and real variable theory. Much detail has been incorporated for this purpose. It is hoped that this book will provide the background necessary to understand what is now happening in probability theory and axiomatic potential theory . . . A glance at the bibliography will reveal an almost total absence of references to books or papers on complex variable type potential theory. This book is primarily about potential theory for higher dimensional spaces, but the two dimensional case is included for the sake of completeness." (In any case the complex variable theory can be found in *Potential Theory on Modern Function Theory*, by M. Tsuji.) As we have indicated, the author's hope is completely satisfied and he has succeeded in making the book accessible to anyone with a background of the basic courses mentioned. If at any time material not found in such courses is needed, it is developed; thus the theory of differentiation of measures is given in the chapter on boundary limit theorems. It is not impossible to read this book with profit without these courses but it would be hard going in places; but the properly prepared reader will have no trouble as the author has a very clear style.

For the reader, prepared as mentioned, who is not a mathematician—and there should be many—it is a pity that the chapter of preliminaries is so short. It could with advantage have contained statements of the main results used—such as the dominated convergence theorem—and have given a few more definitions, such as that of a complete measure, for example. Also the basic courses in real variable should contain a certain amount of general topology for a complete understanding of the chapters on the fine topology and the Martin boundary.

Finally it should be repeated that this book is very well conceived and written and can be recommended as a text for a graduate course and for reading by the interested mathematician or nonmathematician with the appropriate background.

P. S. BULLEN, University of British Columbia

Mathematical Concepts of Elementary Measurement. By A. L. Blakers. SMSG Studies in Mathematics 17. Vromans, Pasadena, California, 1968. 420 pp. $4.00 (paper). (Telegraphic Review, December 1969.)

As announced in the title, the subject is mathematical concepts of measure-

ment rather than measurement in its realistic setting. The allusions to measurement as such are rare and the analysis centers around scalar magnitudes that enjoy a firm foothold in mathematics itself, such as number, length, area, volume, angle. Though vector magnitudes are lacking, even such an important magnitude as orientation is neglected, and the rich variety of concepts of angles is not fully accounted for, the author gives a wealth of information on measurement inside mathematics. Only in the last chapter does the author step outside mathematics to look for measurement ideas with a mathematical background. It is the most valuable part of the work and maybe even its most creative, though it is possible that the same ideas have been conceived independently by many other people or even been published somewhere. It deals with what physicists call dimension theory of physical magnitude. The author recognized that such products as mass times length can be explained as tensor products and that reciprocal magnitudes such as time and frequency can be understood as duals of each other. I wonder why the author did not arrive at interpreting gauge notations as function symbols (of functions from the reals to a magnitude).

The exposition is broad, meticulous and abstract. It is one of the most striking specimens of a philosophy on writing mathematical texts that can be characterized as follows: As to the rather elementary subject matter the reader is expected to know nothing at all, whereas a tremendous skill in the mathematical methods of abstractive thinking is required of him. I guess that the people who have the skill are too well acquainted with the subject matter to read the book, whereas for the others who lack the skill the book is much too difficult. In this respect the most valuable last chapter is the worst. I do not believe that anybody who has not found its content independently would discover what the chapter is about.

HANS FREUDENTHAL, Utrecht, Netherlands

Geometric Inequalities. By O. Bottema, R. Ž. Djordjević, R. R. Janić, D. S. Mitrinović, P. M. Vasić. Wolters—Noordhoff, Groningen, 1969. 151 pp. $4.90. (Telegraphic Review, February 1970.)

Approximately five hundred elementary inequalities for triangles are presented. Proofs, outlines thereof, or references thereto are given for each. Nearly all the inequalities deal with triangles, and all deal with polygons or circles. Algebraic methods of proof are most prominent, and (shades of Newton!) there are no illustrations.

A considerable portion of the book is a translation of the Serbo-Croatian work of the same title by the Yugoslav authors (No. 31, *Matematicka Biblioteka*, Beograd, 1966). The collection contains results published before 1967. There are innumerable elegant results contained in this fascinating collection. High school and college teachers of mathematics alike should find much in it to stimulate their students and themselves.

N. D. KAZARINOFF, University of Michigan

TELEGRAPHIC REVIEWS

Telegraphic reviews are intended to give prompt notice of books, with sufficient information to assist in deciding whether to order an examination copy or suggest library purchase. Possible uses are indicated as follows:

B = college bookstore stock L = library purchase
P = professional reading S = supplementary reading
T = textbook E = teacher education
13 to 18 = freshman to second year graduate level usage
1 to 4 = approximate time in semesters to cover text
* = positive emphasis ? = negative emphasis

Books on high school material (pre-calculus) are denoted REMEDIAL, and normally receive telegraphic reviews only if they are written for college students. Publishers are denoted by the standard abbreviations used in *Books in Print*, which gives complete addresses.

ALGEBRA AND NUMBER THEORY, P, L(RESEARCH). *Arithmetical Functions.* K. Chandrasekharan. Springer-Verlag, 1970, 231 pp, $16. A treatise on the asymptotic behavior of certain arithmetical functions--the partition function, the divisor function, and some related to the distribution of primes. Many of the analytic tools needed are developed in the book, but the author's claim that "only a modicum of acquaintance with analysis and number theory" is presupposed is perhaps a bit optimistic. No problems except unsolved ones. J.D.-B.

ALGEBRA, GROUPS, T(14-15), L. *Representations of Groups: With Special Consideration for the Needs of Modern Physics.* Hermann Boerner. North-Holland, 1970, 341 pp, $16. This is the second edition of this book and includes some subjects not treated in the first edition. The topics are selected with the physicist in mind although the treatment is purely mathematical. R.J.

ALGEBRA, LINEAR ALGEBRA, T(14: 1), B. *Linear Algebra for Social Sciences.* Menahem E. Yaari. P-H, 1971, 174 pp, $9.95. Only the title indicates that this book, written by an economist, was intended for students in the social sciences. There are no applications in the text or problems and there is no bibliography. The material is that in a moderately advanced course in matrix theory plus some set theory, functions, real numbers, and linear inequalities. If you already know applications, this is a very nice book. If not, perhaps you will want to keep looking. W.C.R.

ALGEBRAIC GEOMETRY, P, L. *Abelian Varieties.* David Mumford. *Tata Institute of Fundamental Research Studies in Mathematics.* Oxford U Pr, 1970, 242 pp, $11.75. Lectures given 1967-68 at the Tata Institute; basic analytic theory, algebraic theory in terms of varieties as well as schemes. L.A.S.

ANALYSIS, P, L. *Eleven Papers in Analysis. American Mathematical Society Translations, Series 2, Volume 95.* AMS, 1970, 252 pp, $12.80. Papers translated from Russian on differential equations, operator theory, and Fourier analysis, with a short 1965 article by D.F. Davidenko on the approximate calculuation of determinants. Other authors represented, and the years their papers were published: E.E. Viktorovskiĭ (1965), M.I. Minkevič (1948), K.S. Mamiĭ (1965), F.Z. Ziatdinov (1964, 1965), V.M. Adamjan and

D.Z. Arov (1966), M.I. Višik and G.I. Eskin (1966), V.A. Marčenko (1950), S.B. Stečkin and P.L. Ul'janov (1962), and M.G. Kreĭn (1967). D.F.A.

ANALYSIS, FUNCTIONAL ANALYSIS, P, L. *Functional Analysis and Related Fields; Proceedings of a Conference in honor of Professor Marshall Stone, held at the University of Chicago, May 1968.* Ed: Felix E. Browder. Springer-Verlag, 1970, 241 pp, $16. Eleven research papers and an article by Saunders MacLane on Stone's influence on the origins of category theory, with remarks by Professor Stone following. The research papers include ones by Harish-Chandra, Hewitt, Mackey, Nachbin, Nelson, Segal, Weil, Zygmund, and the editor, and joint ones by Kato and Kuroda and by Chern, do Carmo and Kobayashi. D.F.A.

DIFFERENTIAL AND INTEGRAL CALCULUS, T(13). *Calculus: One and Several Variables.* Saturnino L. Salas and Einar Hille. Xerox Coll, 1971, 800 pp, $13.95. This book is designed for a three-semester course in calculus. The information in the first two-thirds of the book is essentially the same as that found in *First-Year Calculus* by Hille and Salas, with sections on the differential and implicit differentiation now included. For those who found the authors' *First-Year Calculus* a stiff exposition from the beginning calculus student's point of view, *Calculus* is somewhat softened, but could not be called intuitive calculus. R.J.

DIFFERENTIAL AND INTEGRAL EQUATIONS, MATHEMATICAL PHYSICS, T(17-18: 2), L. *Mathematical Physics, An Advanced Course: Applied Mathematics and Mechanics, Volume II.* S.G. Mikhlin. North-Holland, 1970, 561 pp, $30. An advanced work (functional analysis level) in mathematical physics which primarily treats linear partial differential equations. The direction of the book is well-conceived and the subjects taken are substantial. This would be an excellent book to work through. R.J.

DIFFERENTIAL AND INTEGRAL EQUATIONS, PARTIAL DIFFERENTIAL EQUATIONS, T(16-17: 1), L. *Partial Differential Equations: Applied Mathematical Sciences, Volume I.* F. John. Springer-Verlag, 1971, 221 pp, $6.50 (P). This book could well be used as a text or supplementary text for a first course in partial differential equations. An up-to-date list of books for further study in this field is given. R.J.

DIFFERENTIAL AND INTEGRAL EQUATIONS, SEVERAL VARIABLES, P, L. *Lecture Notes in Mathematics, Volume 167: Multidimensional Inverse Problems for Differential Equations.* M. Lavrentiev, V. Romanov and V. Vasiliev. Springer-Verlag, 1970, 59 pp, $2.80 (P). A monograph which investigates a number of multidimensional inverse problems, giving exposure to new approaches for solutions. R.J.

DIFFERENTIAL AND INTEGRAL EQUATIONS, TOPOLOGY, T*(16-18: 1, 2), L. *Stability Theory of Dynamical Systems.* N.P. Bhatia and G.P. Szegö. *Die Grundlehren der Mathematischen Wissenschaften, Band 161.* Springer-Verlag, 1970, 225 pp, $16. An introduction to the theory of dynamical systems in metric spaces, with emphasis on stability theory and its applications to autonomous systems of ordinary differential equations. This book is very welcome indeed, as it blends standard material with results of recent research in a text

designed for strong undergraduates and beginning graduate students; the only prerequisites are elementary courses in ordinary differential equations, analysis, and topology. Many solid, theoretical exercises and over 750 references; applications to ordinary differential equations include a study of Liapunov functions. D.F.A.

DIFFERENTIAL EQUATIONS, T??, S, L. *Differential Equations and Applications for Students of Mathematics, Physics, and Engineering.* James B. Scarborough. Robert E. Krieger, 1965, 479 pp. Contains a brief (131 pp) cookbook for the solution of differential equations common to a first course, but the major portion of the book is devoted to applications, varying from simple exponential decay to the problem of the silent bell. The first part contains very few exercises, the second practically none. Has definite value as a source for examples, very questionable value as a text. T.A.V.

ECONOMICS, P. *Theory of Cost and Production Functions.* Ronald W. Shephard. Princeton U Pr, 1970, 308 pp, $15. A revision, at the urging of Oskar Morgenstern, of the author's 1953 monograph *Cost and Production Functions.* The production function is defined with respect to an axiomatically defined "production technology", and is not subject to special conditions which are apparently traditional. R.B.K.

FINITE SECTIONS OF INEQUALITIES, P, L. *Finite Sections of Some Classical Inequalities.* Herbert S. Wilf. *Ergebnisse der Mathematik und Ihrer Grenzgebiete, Band 52.* Springer-Verlag, 1970, 83 pp, $7.70. A treatise for the serious mathematician working in the area given by the title of the book. This work brings together lines of research which focus on the spectral theory of finite-dimensional sections of infinite-dimensional operators. Central in the development is the question of the rate at which the spectra of sections converge to the spectrum of the operators. R.J.

FOUNDATIONS, AUTOMATA THEORY, CATEGORY THEORY, L, P. *Recursiveness.* Samuel Eilenberg and Calvin Elgot. Acad Pr, 1970, 89 pp, $6.50. Category theory of recursive functions, algebraic statements of arithmetic theorems. Introduction to categorical notation, recursive and primitive recursive functions and relations, R.E. sets. Despite publisher's claims, some skill in arrow-chasing is a prerequisite. J.G.L.

FUNCTIONAL ANALYSIS, P, L. *Analyse non-archimédienne.* A.F. Monna. *Ergebnisse der Mathematik und Ihrer Grenzgebiete, Band 56.* Springer-Verlag, 1970, 119 pp, $10.50. Analysis done over fields with a non-Archimedean valuation (e.g. p-adic fields), series, topological vector spaces, Banach spaces, convexity, duality, integration--all non-Archimedean. Large bibliography. L.A.S.

GALOIS THEORY, P, L. *Galoissche Theorie der p-Erweiterungen.* H. Koch. Springer-Verlag, 1970, 161 pp, $12. A unified treatment of results of Šafarevič, Fröhlich, Brumer and the author, on the Galois theory of p-extensions. The development is based largely on the cohomology theory of profinite groups, treated in the first chapter. J.D.-B.

GENERAL, P, L. *Eleven Papers on Logic, Algebra, Analysis and Topology. American Mathematical Society Translations, Series 2,*

Volume 97. AMS, 1971, 258 pp, \$13.70. Translations of articles in Russian on (with date of publication): logic, by S. Jr. Maslov (1964); group theory, by O.N. Macedonskaja (1966); functions of a complex variable, by E.P. Dolženko (1963) and M.Š. Stavskiĭ (1966); operator theory, by I.C. Gohberg (1966), M.G. Kreĭn (1949), and Ju. L. Šmul'jan (1968); dynamical systems, by B.M. Budak (1952) and I.V. Bronšteĭn (1963); and mappings of Euclidean space R^n into itself and on the real line into itself, by A.N. Šarkovskiĭ (1965 and 1966, respectively). D.F.A.

GEOMETRY, P, S, L. *Lecture Notes in Mathematics, Volume 158: Finite Translation Planes*. T.G. Ostrom. Springer-Verlag, 1970, 122 pp, \$2.80 (P). A series of lectures presented as two independent but complementary books. Book I shows ways in which replaceable nets can be used to construct previously unknown planes. Book II uses the theory of linear groups in an attempt to classify planes according to the central collineations which they admit. J.N.C.

GEOMETRY, ALGEBRAIC GEOMETRY, P, L(RESEARCH). *Lecture Notes in Mathematics, Volume 176: Fundamentalgruppen algebraischer Mannigfaltigkeiten*. Herbert Popp. Springer-Verlag, 1970, 156 pp, \$4.40 (P). An expansion of the notes for lectures given by the author at Heidelberg and the University of British Columbia, surveying the beginnings which have been made in obtaining by algebraic methods an understanding of Zariski's work on the fundamental groups of algebraic varieties. J.D.-B.

GEOMETRY, ALGEBRAIC GEOMETRY, P, L(RESEARCH). *Lecture Notes in Mathematics, Volume 174: Linear Determinants with Applications to the Picard Scheme of a Family of Algebraic Curves*. Birger Iversen. Springer-Verlag, 1970, 69 pp, \$2.20 (P). Generalizes Weil's construction of the Jacobian variety of an algebraic curve "to construct Picard schemes, in the sense of Grothendieck, for a flat and proper family of geometrically reduced and irreducible curves.J.D.-B.

GEOMETRY, ANALYTIC, T(13: 1), L. *Analytic Geometry*. A.C. Burdette. Acad Pr, 1971, 225 pp, \$6.50. An attractive presentation which allows considerable flexibility in the depth of treatment. A few applications of conic sections are included, but the book would create more interest if the number and diversity of applications were increased. J.N.C.

GEOMETRY, RELATIVITY, ASTRONOMY, PHILOSOPHY, T(18), P, L. *Relativity and the Question of Discretization in Astronomy*. Dominic G.B. Edelen and Albert G. Wilson. *Springer Tracts in Natural Philosophy, Volume 20*. Springer-Verlag, 1970, 186 pp, \$10.50. The purpose of this highly technical work is to shed light on the question of whether the structure of the physical world is more adequately described by a continuous or by a discrete (quantum) mode of representation. Those who have not worked in this special area will find much of the terminology out of their world. R.J.

HISTORY. *Des Révolutions des Orbes Célestes*. Nicolas Copernic. Transl: A. Koyré. A. Blanchard, 1970, 154 pp, \$3.30. A new edition, with notes and errata, of the Latin text together with French translation. L.A.S.

HISTORY, COMPUTER SCIENCE, FOUNDATIONS, L*, S. *Cybernetics for the Modern Mind.* Walter R. Fuchs. Transl: K. Kellner. Macmillan, 1971, 357 pp, $6.95. A philosophical and historical survey of the major results of modern logic, linguistics and computer science. Despite the publisher's extravagant claims about "Electronic Brains", this is a literate and interesting popularization of many of the current applications of linguistics. Sometimes long-winded (perhaps in translation). No bibliography. J.G.L.

HISTORY, MODERN, L. *Colloques Textes Des Rapports.* Actes du XII^{e} Congrès International d'Histoire des Sciences, Tome I(A). Albert Blanchard, 1970, 431 pp, $9.30. One of the colloquia, included in this first of 13 volumes reporting on the XII^{e} Congrès, in Paris in 1968, consists of five papers (in French) on the origins of modern algebra. L.A.S.

HISTORY, PHILOSOPHY, GEOMETRODYNAMICS, T(14-15), B, L. *The Conceptual Foundations of Contemporary Relativity Theory.* John Cowperthwaite Graves. MIT, 1971, 361 pp, $15. A historical but primarily philosophical treatment of the general theory of relativity with exposure of recent developments in the theory of geometrodynamics. R.J.

ISOMETRIC EMBEDDINGS, P, L. *Isometric Embeddings of Riemannian and Pseudo-Riemannian Manifolds.* Robert E. Greene. *Memoirs of the American Mathematical Society, Number 97.* AMS, 1970, 63 pp, $1.80 (P). Nash's results are generalized to isometric embeddings for arbitrary quadratic forms on the tangent bundle of a manifold. The essential tool is an implicit function theorem of Schwarz which is proved in detail using the Schauder fixed point theorem and Nash's smoothing operators. A local C^∞ theory then shows that any pseudo-Riemannian manifold is embeddable isometrically in any other pseudo-Riemannian manifold of sufficiently high dimension and appropriate signature. R.B.K.

JORDAN ALGEBRAS, P, L. *Octonion Planes Defined by Quadratic Jordan Algebras.* John R. Faulkner. *Memoirs of the American Mathematical Society, Number 104.* AMS, 1970, 71 pp, $1.50 (P). A doctoral dissertation written under Jacobson. "Results on octonion (Cayley, octave, Moufang) planes are derived in a uniform fashion...by using an exceptional quadratic Jordan algebra." Some new results appear, such as the case of characteristic two. No index. W.C.R.

OPERATOR THEORY, FUNCTIONAL ANALYSIS, P, L. *Proceedings of Symposia in Pure Mathematics, Volume XVIII, Part I: Nonlinear Functional Analysis.* AMS, 1970, 296 pp, $12. Twenty-two papers given at the AMS Symposium on Nonlinear Functional Analysis held in Chicago in April, 1968. Authors are Asplund, Berger, Brezis, Cesari, Cronin, Eells, Elworthy and Tromba, de Figueiredo and Karlovitz, Fujita, Halpern, Jones, Kato, Kirk, Lions, Mosco, Palais, Petryshyn, Rabinowitz, Rockafeller, Rothe, Stampacchia, Strauss. D.F.A.

OPTIMIZATION AND LINEAR PROGRAMMING, T*(16-17: 1), S, P, L. *Linear Optimization.* W. Allen Spivey and Robert M. Thrall. HR & W, 1970, 530 pp, $15.75. A well-organized development from satisfying examples to a clean mathematical theory plus appendices on foundations and linear algebra allow this book to be considered as a text for different purposes and levels. In addition to applications contained in the multitude of examples and exercises, there are chapters on

assignment, transportation, and game theory problems. Several algorithms are described in detail with flow charts appended. R.W.N.

PARTIAL DIFFERENTIAL EQUATIONS, P, L. *Boundary Value Problems of Mathematical Physics. V.* Ed: O.A. Ladyzenskaja. *Proceedings of the Steklov Institute of Mathematics, Number 102 (1967).* AMS, 1970, 185 pp, $16.10 (P). Seven articles by five authors on motions of viscous incompressible fluids, the Dirichlet problem in unbounded domains, a priori estimates for solutions of various problems, and a generalization of Marcinkiewicz's interpolation theorem. Authors are K.K. Golovkin, A.V. Ivanov, A.P. Oskolkov, V.A. Solonnikov, and the editor. D.F.A.

PHILOSOPHY OF SCIENCE, THEORETICAL PHYSICS, P, L(RESEARCH). *Lecture Notes in Physics.* G. Ludwig. Springer-Verlag, 1970, 469 pp, $7.70 (P). Subtitle: *(Deutung des Begriffs physikalische Theorie und axiomatische Arundlegung der Hilbertraumstruktur der Quantenmechanik durch Hauptsatze des Messens). Explanation of the concept 'physical theory' and axiomatic foundation of the Hilbert space structure of quantum mechanics on fundamental principles of measurement.* J.D.-B.

PHYSICS, S, B. *Special Theory of Relativity.* C.W. Kilmister. Pergamon Pr, 1970, 299 pp, $7, $4.75 (P). An introduction to special relativity, designed for undergraduates. The first third of the book, written by Kilmister, discusses the theory, Einstein's contribution to it, its applications in quantum theory, and some elementary consequences of the Lorentz transformation. The remainder is a collection of original writings on topics in special relativity by Michelson, Larmor, Lorentz, Poincaré, Einstein, Wilson and Wilson, Zeeman, Dirac, Anderson, and Wigner. Wigner's papers concerns unitary representations of the inhomogeneous Lorentz group and is part of a 1939 paper which appeared in the *Annals of Mathematics.* D.F.A.

PROBABILITY, T(17-18: 2, 3), S, L. *Probability Theory.* A. Rényi. *Applied Mathematics and Mechanics Series, Volume 10.* North-Holland, 1970, 666 pp, $21.50. This translation (by L. Vekerdi) has some clumsy English in it, but is generally quite readable. Chapter on information theory included. Presupposes the theory of real and complex variables. Extensive problem sets and bibliography. F.L.W.

PROBABILITY AND STATISTICS, T(16: 2). *Probability and Statistics.* Harry Lass and Peter Gottlieb. A-W, 1971, 470 pp, $12.95. This text differs from most in its strong emphasis on combinatorial analysis. The treatment of probability is very thorough, but the statistics portion, in particular hypothesis testing, is slighted. A 91-page appendix covering such topics as difference equations, matrix theory, Jacobians and Lagrange multipliers provides necessary mathematical background. The book also contains an abundance of examples and problems. R.S.K.

PROBABILITY AND STATISTICS, T(17: 2), P. *An Introduction to Probability, Decision, and Inference.* Irving H. LaValle. HR & W, 1970, 767 pp, $11.95. This text gives an introduction to probability and statistics with a strong emphasis on the Bayesian point of view. The topics covered range from the very elementary to the quite complex, resulting in an admitted nonuniformity of style.

The book includes much material on Bayesian decision theory and relates Bayesian and orthodox approaches to statistical inference. However, nothing on regression or analysis of variance is included. R.S.K.

REAL ANALYSIS, T?(14-15: 1). *Introduction to Real Analysis.* Michael Gemignani. Saunders, 1971, 160 pp, $8. A weak version of a typical introductory book, beginning with the development of the reals and ending with uniform convergence, with no features to distinguish it from a host of others. T.A.V.

REAL ANALYSIS, CALCULUS, T(14: 2). *Calculus and Linear Algebra: Combined Edition.* Wilfred Kaplan and Donald J. Lewis. Wiley, 1971, 1159 pp, $14.95. This is a combined edition of Volumes 1 and 2 with the last chapter of Volume 2 (on differential equations) deleted. Telegraphic review of Volume I, August 1970 and of Volume II, May, 1971. L.L.K.

REAL ANALYSIS AND CALCULUS, T(13). *Mathématiques. Eléments de Calcul Différentiel et Intégral. Tome I.* A. Hocquenghem, P. Jaffard and R. Chenon. Masson & Cie, 1971, 552 pp. A substantial text for good students, written in French, for use in a course in differential and integral calculus. R.J.

REAL ANALYSIS, LINEAR ALGEBRA, T(15: 2). *Mathématiques. Algèbre Linéaire, Représentation Des Fonctions Analyse Vectorielle, Equations Fonctionnelles. Tome II.* A. Hocquenghem, P. Jaffard and R. Chenon. Masson & Cie, 1970, 522 pp. A text, written in French, giving a strong introductory treatment of linear and multilinear algebra, representations of functions, vector analysis and differential equations. R.J.

STATISTICS, NONPARAMETRIC, P*, L. *Nonparametric Techniques in Statistical Inference.* Ed: Madan Lal Puri. Cambridge U Pr, 1970, 623 pp, $32.50. This book contains the proceedings of the first International Symposium on Nonparametric Techniques in Statistical Inference, which was held at Indiana University in June, 1969. The goal of the symposium was to provide current information and stimulate further research in this increasingly important area of statistics. Acknowledged experts presented research papers in the areas of testing and estimation, order statistics and allied problems, general theory, ranking and selection procedures, and decision theoretic and empirical Bayes procedures. Also included is one paper on the teaching of nonparametric statistics. R.S.K.

SURGERY, P*, L*. *Surgery on Compact Manifolds.* C.T.C. Wall. Acad Pr, 1971, 280 pp, $14.50. A delightful and thorough exposition of the operation introduced in 1959 by Dr. Milnor. The style is comfortably informal and the exposition together with its extensive references carries the development and applications of surgery (spherical modification, χ-equivalence, Morse reconstruction) up to 1970. J.A.S.

TOPOLOGY, T*(16-17: 2), S, P, B, L. *Topology for Analysis.* Albert Wilansky. Ginn, 1970, 383 pp, $13.50. The author claims this text "begins with first principles and develops without haste all that part of topology which may be described as generalized analysis." The choice of such topics is necessarily arbitrary, but the author

does a very creditable job. Although there seems to be a heavy-handed use of reference to previous results simply by number, the book is very readable and a valuable addition to the literature. In the preface the author warns of the disease of "axiomatics", but succumbs himself, with an extensive appendix of equivalences, implications and counterexamples. T.A.V.

TRANSLATIONS, P. *Fifteen Papers on Real and Complex Functions, Series, Differential and Integral Equations. American Mathematical Society Translations, Series 2, Volume 86.* AMS, 1970, 282 pp, $14.40. Papers on entire functions, decrease of harmonic functions, function classes, differentiability relative to congruent sets, representability of functions (Hilbert's 13th problem), trigonometric series, Carlson's inequality, a Tauberian theorem, convergence of Fourier series, series relating to a system, elliptic equations, methods of Galerkin type, the Čaplygin problem. R.B.K.

TRANSLATIONS, P. *Sixteen Papers on Differential and Difference Equations, Functional Analysis, Games and Control. American Mathematical Society Translations. Series 2, Volume 87.* AMS, 1970, 303 pp, $15.40. Papers on polynomials orthogonal with weight, fundamental solutions invariant under rotation, the Tricomi problem, Petrovskiĭ-correct equations, asymptotic behavior for parabolic equations, differential equations in a Banach space, imbedding theorems, positive linear operators, expansions in eigenvectors and associated vectors, differential operators of infinite order, nonlinear systems of difference equations, automata games, extrapolation in automatic control. R.B.K.

VECTOR ANALYSIS, T(14), S, B. *Vector Analysis.* L. Marder. Am Elsevier, 1970, 167 pp, $7.50. Traditional vector analysis in 2 and 3 dimensions. "The emphasis is on motivation by physical illustrations..." Neither general vector spaces nor tensors is discussed. Contains a reasonable number of exercises with answers and an appendix on differentials. T.A.V.

VECTORS, T(13-14: 1), S, L. *Introduction to Vector Analysis.* J.C. Tallack. Cambridge U Pr, 1970, 298 pp, $4.75. Primarily of interest to students of science and engineering who wish for an easy introduction to the fundamental principles of vector analysis. Six chapters have been added to the author's *Introduction to Elementary Vector Analysis.* These cover: (i) new techniques (vector product and triple products) and (ii) applications. L.C.L.

Reviewers Whose Initials Appear Above

David F. Appleyard, Carleton; Judith N. Cederberg, St. Olaf; John Dyer-Bennet, Carleton; Richard Jarvinen, Carleton; Lorraine L. Keller, St. Olaf; Roger B. Kirchner, Carleton; Richard S. Kleber, St. Olaf; Loren C. Larson, St. Olaf; John G. Lewis, St. Olaf; R.W. Nau, Carleton; William C. Ramaley, Carleton; J. Arthur Seebach, Jr., St. Olaf; Linda A. Seebach, St. Olaf; T.A. Vessey, St. Olaf; Frank L. Wolf, Carleton.

NEWS AND NOTICES

EDITED BY RAOUL HAILPERN, SUNY at Buffalo

Readers are invited to contribute to the general interest of this Department by sending news items to Mathematical Association of America, 1225 Connecticut Avenue, NW, Washington, D. C. 20036. Items must be submitted at least two months before publication can take place.

PERSONAL ITEMS

Professor H. M. Gehman, SUNY at Buffalo, represented the Association at the inauguration of Dr. R. L. Ketter as President of SUNY at Buffalo on February 15, 1971.

Bowling Green State University: Drs. J. K. Brown, Michigan State University, and Victor Norton, University of Michigan, have been appointed Assistant Professors; Associate Professor W. L. Terwilliger has been appointed Assistant Chairman of the Department of Mathematics.

Eastern Nazarene College: Dr. Floyd John, Raytheon, has been appointed Professor; Associate Professor G. E. Lashley has been appointed Chairman of the Department of Mathematics.

Georgia Institute of Technology: Professor B. M. Drucker has resigned as Director of the School of Mathematics and has returned to teaching; Associate Professor C. H. Holton retired on December 31, 1970, after thirty-three years at Georgia Institute of Technology; Associate Professor J. D. Neff has been appointed Acting Director of the School of Mathematics; Dr. H. S. Valk, University of Nebraska, has been appointed Dean of the General College.

Louisiana State University, New Orleans: Professor J. R. Foote, University of Missouri, Rolla, has been appointed Professor; Assistant Professor C. S. Rees, University of Tennessee, has been appointed Assistant Professor; Assistant Professor C. L. Outlaw has been promoted to Associate Professor.

Miami University: Associate Professor R. G. Laatsch has been promoted to Professor; Assistant Professor S. E. Payne has been promoted to Associate Professor.

Naval Postgraduate School: Drs. Craig Comstock, University of Michigan, and Peter C. C. Wang, State University of Iowa, have been appointed Associate Professors; Drs. W. C. Chewning, University of Virginia, and G. A. Stoops, Litton Industries, have been appointed Assistant Professors.

Wisconsin State University, La Crosse: Drs. J. M. Sobota, Michigan State University, and J. D. Wine, Virginia Polytechnic Institute, have been appointed Assistant Professors.

Dr. P. M. Bailyn, The Cooper Union, has been promoted to Associate Professor.

Mr. Carl Baker, Southern Connecticut State College, has been appointed Assistant Professor at New England College.

Professor W. T. Graybeal, Emory and Henry College, has been appointed Assistant Academic Dean and Registrar.

Dr. G. E. Hedrick, Iowa State University and USAEC, has been appointed Assistant Professor, Department of Computing and Information Sciences, Oklahoma State University.

Professor Louis Nirenberg, New York University, has been named Director of the University's Courant Institute of Mathematical Sciences.

Professor L. V. Quintas, Pace College, has been appointed Chairman of the Mathematics Department.

Professor A. F. Strehler, Carnegie-Mellon University, has been appointed Dean of Graduate Studies.

Assistant Professor P. R. Thie, Boston College, has been promoted to Associate Professor.

Assistant Professor Steven Thomason, Simon Fraser University, has been promoted to Associate Professor.

Professor Nura Turner, SUNY at Albany, retired on July 1, 1970 with the title of Professor Emeritus.

Assistant Professor J. H. M. Whitfield, Lakehead University, has been promoted to Associate Professor.

Associate Professor E. W. Bailey, Indiana University of Pennsylvania, died on June 30, 1970 at the age of 58. He was a member of the Association for five years.

Emeritus Professor Edith A. McDougle, University of Delaware, died on September 8, 1970 at the age of 78. She was a member of the Association for forty years.

Professor J. W. T. Youngs, University of California, Santa Cruz, died on July 20, 1970 at the age of 59. He was a member of the Association for twenty-four years.

NATIONAL MEDAL OF SCIENCE WINNERS FOR 1970

In January the White House announced the National Medal of Science winners for 1970. The award, established in 1959 and presented annually, is the highest award of the federal government for distinguished achievement in science, mathematics, and engineering. One of the recipients is Professor Richard D. Brauer, of Harvard University, honoured for his development of the theory of modular representations.

PUBLICATIONS OF THE CANADIAN MATHEMATICAL CONGRESS

The Canadian Mathematical Congress announces the following publications, which may be purchased from its Montreal Office. The address is 985 Sherbrooke West, Montreal 110, Canada. Note the discount available to members of the Association.

Proceedings of the Twelfth Biennial Seminar on Time Series Stochastic Processes, Convexity, Combinatorics, held at the University of British Columbia, August 11–27, 1969. Edited by R. Pyke. xiv+304 pages. Cdn. $20 plus postage charges of $1.00. (20% discount to any member of a recognized mathematical society.)

Counting Labelled Trees, by J. W. Moon. x+113 pages. Cdn. $5 plus postage charges of 50¢. (20% discount to any member of a recognized mathematical society.)

Introduction to Markov Chains, by D. A. Dawson. 103 pages. Cdn. $3 plus postage charges of 50¢.

THE ROCKY MOUNTAIN JOURNAL OF MATHEMATICS

The Rocky Mountain Mathematics Consortium announces the publication of THE ROCKY MOUNTAIN JOURNAL OF MATHEMATICS to begin in 1971. The journal will be issued quarterly and will publish both primary research and survey articles. The first issue of THE ROCKY MOUNTAIN JOURNAL OF MATHEMATICS will consist of a series of lectures from an Advanced Science Seminar held at Northern Arizona University in Flagstaff on Mathematical Theory of Scattering presented under the auspices of The Rocky Mountain Mathematics Consortium and supported by a grant from the National Science Foundation.

On the Editorial Board are Professors William Scott, Managing Editor, University of Utah, Harvey Cohn, University of Arizona, Bernard Epstein, University of New Mexico, Lloyd K. Jackson, University of Nebraska, and Wolfgang Thron, University of Colorado. Papers may be submitted to any of the editors. Inquiries may be sent to Professor Robert W. McKelvey, Executive Director of The Rocky Mountain Mathematics Consortium, Mathematics Building, University of Montana, Missoula, Montana 59801.

MATHEMATICAL ASSOCIATION OF AMERICA

Official Reports and Communications

OCTOBER MEETING OF THE INDIANA SECTION

The fall meeting of the Indiana Section of the MAA was held on October 24, 1970, at Wabash College, Crawfordsville. About 70 persons attended, of whom 65 were members of the Association.

A welcome was extended by the Chairman of the Section, Professor W. C. Swift. The morning session consisted of an address by Professor R. E. Zink, Purdue University, "On the Representation of Measurable Functions by Series." This was followed by a presentation and discussion of the question of accreditation of departments of mathematics and certification of graduates, led by Professor B. E. Rhoades, Indiana University, and Professor R. G. Bartle, University of Illinois. Despite the efforts at impartiality on the part of the panel, there was no enthusiasm for the proposal among those present. One letter from an absent member supporting the proposal was read.

At the afternoon business meeting the Chairman announced his intention to appoint a committee to revise the by-laws of the Section. Following the business meeting, Professor W. P. Ziemer, Indiana University, gave an address entitled, "Some Recent Developments in the Plateau Problem."

R. T. Hood, *Secretary-Treasurer*

NOVEMBER MEETING OF THE NORTH CENTRAL SECTION

The annual fall meeting of the North Central Section of the MAA was held at the North Dakota State University, Fargo, on November 7, 1970. Professors Robert Tidd and Charles Friese, of North Dakota State University, presided at the morning sessions and Professor Alfred Aeppli, University of Minnesota, presided at the afternoon session. One hundred thirty persons attended, including one hundred three members.

Professor I. N. Herstein, University of Chicago, gave the invited address: "Mappings Related to Homomorphisms."

Other papers presented were:

1. *Computer Supplemented Finite Algebra*, by J. F. Peters, Saint John's University, Collegeville, Minnesota.
2. *A Determinant for the Hermite Polynomial $H_n(x)$*, by F. J. Arena, North Dakota State University, Fargo, North Dakota.
3. *Distributions whose Test Functions are Sequences*, by Clayton Knoshaug, Bemidji State College, Bemidji, Minnesota.
4. *Dirichlet Series Obtained by Iteration*, by George Brauer, University of Minnesota, Minneapolis, Minnesota.
5. *Convex Functions and Differential Inequalities*, by R. M. Mathsen, North Dakota State University, Fargo, North Dakota.
6. *The Distribution of Quadratic Residues in Fields of Order p^2*, by G. E. Bergum, South Dakota State University, Brookings, South Dakota.
7. *If T is a Torsion Group, HOM (T, G) is Algebraically Compact*, by Milton Legg, Moorhead State College, Moorhead, Minnesota.
8. *Algebraic Closure: a Non-Standard Approach*, by L. C. Larson, St. Olaf College, Northfield, Minnesota.
9. *Factoring Functions on Cartesian Products*, by Milton Ulmer, Macalester College, St. Paul, Minnesota.
10. *Rational Numbers Generated by Two Integers*, by G. A. Heuer, Concordia College, Moorhead, Minnesota.

11. *Panel Discussion—The Question of Accreditation and Certification.* Moderator: Alfred Aeppli, University of Minnesota. Panelists: Warren Loud, University of Minnesota; Robert Earles, St. Cloud State College; Murray Braden, Macalester College.

W. J. THOMSEN, *Secretary-Treasurer*

NOVEMBER MEETING OF THE NORTHEASTERN SECTION

The sixteenth annual meeting of the Northeastern Section of the MAA was held at Merrimack College, North Andover, Massachusetts, on November 28, 1970. The Section Chairman, M. C. Gemignani of Smith College, presided. One hundred and sixty-four people attended the meeting, of whom 134 were members of the Association.

The morning meeting was devoted to a panel discussion: "Calculus and Introductory Mathematics," with M. C. Gemignani, Smith College, moderator, and the following panelists: D. E. Christie, Bowdoin College; N. R. Grabois, Williams College; A. P. Mattuck, M.I.T.

At the business meeting the following officers were elected: Chairman, R. D. Schafer, Massachusetts Institute of Technology; Vice Chairman, D. L. Kreider, Dartmouth College; Secretary-Treasurer, G. W. Best, Phillips Academy. The newly elected Vice Chairman, D. L. Kreider, was appointed chairman of a committee to investigate the possible implementation of a High School lecture program. The business meeting concluded with Professor Grace E. Bates chairing a discussion of the CUPM Report on Accreditation and Certification.

The afternoon program consisted of three panel discussions:

Junior College Mathematics. Moderator: J. A. Fickes, Quinnipiac College. Panel: Grace E. Bates, Mount Holyoke College; Oscar Sassian, Holyoke Junior College.

Teacher Training. Moderator: Rev. S. J. Bezuszka, S. J., Boston College. Panel: D. L. Kreider, Dartmouth College; B. E. Meserve, University of Vermont.

Preparation of Research Mathematicians. Moderator: N. H. McCoy, Smith College. Panel: Haskell Cohen, University of Massachusetts; M. E. Munroe, University of New Hampshire.

G. W. BEST, *Secretary-Treasurer*

NOVEMBER MEETING OF THE OHIO SECTION

A special meeting of the Ohio Section of the MAA was held on November 7, 1970, at Ohio Northern University, Ada, Ohio. The subject of the meeting was accreditation and certification. Eighty-two persons were registered in attendance, including seventy members of the Association. Professor B. J. Yozwiak, Chairman of the Section, and Professor R. G. Laatsch, Chairman of the Program Committee, presided. The following program was presented:

1. *The American Chemical Society accreditation and certification program,* by Donald Bettinger, Chairman of the Department of Chemistry, Ohio Northern University.
2. *The CUPM report on accreditation and certification,* by D. T. Finkbeiner II, Kenyon College.
3. *Panel discussion.* Panel members: Harold Brown, The Ohio State University; Louis Graue, Bowling Green State University; James Murtha, Marietta College; and Donald Peter, Lorain County Community College.

FOSTER BROOKS, *Secretary-Treasurer*

CONTRIBUTIONS FROM MAA MEMBERS, 1970 AND 1971

In August, 1969 the Board of Governors voted to create three new types of membership to recognize members who wished to contribute to the support of the Association beyond the level of normal dues. We are pleased to announce that in 1970 and 1971 a total of 163 persons were designated Contributing Members in recognition of voluntary dues payments of $25.00.

A member is designated a Sponsor in any year in which he pays dues of between $50.00 and $99.00 and a Patron in any year in which he pays dues of at least $100.00. The Association gratefully honors its Sponsors and Patrons for 1970 and 1971, who are listed here:

SPONSORS

1970	*1971*
Anonymous	A. B. Cunningham
A. B. Cunningham	W. E. Hartnett
L. S. Dederick	Victor Klee
W. R. Harris, Jr.	Mrs. Alda F. Oertly
W. E. Hartnett	I. J. Schoenberg
S. A. Jennings	Mrs. Elizabeth B. Weinstock
Victor Klee	
I. J. Schoenberg	

PATRONS

Anonvmous	Anonymous
C. B. Allendoerfer	C. B. Allendoerfer
F. H. Dekoven	Leonard Gillman
Leonard Gillman	W. S. Lewis
H. A. Gindler	

We also gratefully acknowledge a contribution of $100.00 from the Addison-Wesley Publishing Company.

A Life Member is designated a Patron Life Member when he pays life dues of $300.00. As of April 1, 1971, the Patron Life Members of the Association are:

N. H. Ball	Mina S. Rees
H. M. Gehman	A. W. Tucker
G. B. Price	R. J. Walker

CUPM CONSULTANTS BUREAU

The members of the Consultants Bureau are some thirty-nine mathematicians chosen for their experience, educational interests, professional specialities, and geographical location. They are available for two-day visits to colleges, junior colleges, and state departments of education.

During his visit a consultant is prepared to discuss questions of all kinds pertaining to undergraduate mathematics curricula. CUPM has made many recommendations concerning curricula; its consultants are willing to discuss them and to offer suggestions for their implementation or adaptation to local needs. A consultant may also discuss problems connected with staffing the mathematics department: the question of the educational qualifications of the staff, ways to strengthen the department, and sources of funds for faculty fellowships and for other programs designed to improve undergraduate instruction.

In some cases, more extensive participation by the consultant is required. To meet these needs CUPM has inaugurated an extended consultant service which will involve a series of visits by the consultant to the school, perhaps every two weeks during the course of the semester. The duties of such a consultant will be determined to some extent on an *ad hoc* basis, but the experience of the Consultants Bureau suggests the two following possibilities which might, of course, be combined:

(a) A department which plans to introduce a course that differs significantly from those previously taught by its faculty might seek a qualified consultant to lead

an intensive faculty seminar covering the course material, as well as the pedagogical difficulties which might be expected in the course. Example: The CUPM recommended course in multivariate calculus using differential forms.

(b) A department which is planning a thorough curriculum revision might seek a consultant to act essentially as a member of its curriculum committee, taking part in its discussions of courses, texts, staffing problems, library, etc.

Those who wish one or more visits by a CUPM consultant should not be deterred by lack of funds; an applicant's financial ability to support the program is not a factor in approving applications. For details of the programs and an application form, write to CUPM, Box 1024, Berkeley, California, 94701.

CUPM REPORT: A COURSE IN BASIC MATHEMATICS FOR COLLEGES

In January, 1970, the Committee on the Undergraduate Program in Mathematics (CUPM) initiated a study concerning curricular problems for those students who are deficient in basic mathematics. There is a sizeable number of college students enrolled in mathematics courses below the level of college algebra and trigonometry, and it is CUPM's belief many of these students can be greatly helped by a reform in this lower level of the mathematics curriculum. Accordingly, CUPM has now prepared a report, *A Course in Basic Mathematics for Colleges* (referred to as Mathematics E).

It is proposed that some of the currently existing basic mathematics courses be replaced by this flexible one-year course, together with an accompanying mathematics laboratory. The laboratory would serve to remedy the students' arithmetic deficiencies, offer added opportunity for drill in algebraic manipulations and allow for instruction in several vocational-oriented topics. The main aim of this course will be to provide the students with enough mathematical literacy for adequate participation in the daily life of our present society.

Many of the students in standard basic mathematics courses have seen the same material in elementary and secondary schools, and it is often the case that this second exposure is no more successful than the first. Thus, a new and more appropriate approach is needed to meet the needs of college students.

In Mathematics E it is recommended that flow-charting and algorithmic and computer-related ideas be introduced early and used throughout. This should give the student a technique in the analysis of problems and encourage him to be precise in dealing with both arithmetic and non-arithmetic operations. Topics of everyday concern, such as how bills are prepared by a computer, calculation of interest in installment buying, quick estimation, analyses of statistics appearing in the press, and various job-related algebraic and geometric problems, are mainstays of the course.

In order to make the recommendations as clear as possible, a topical outline, with an extensive commentary, is given. However, the outline should be viewed more as a flexible model rather than a rigid description; the spirit of the course is more important than content. The model outline contains flow charts and elementary operations, rational numbers, geometry I, linear polynomials and equations, the computer, nonlinear relationships, geometry II, statistics, and probability.

A Course in Basic Mathematics for Colleges is available without charge from CUPM, P.O. Box 1024, Berkeley, California 94701.

CALENDAR OF FUTURE MEETINGS

Fifty-second Summer Meeting, Pennsylvania State University, University Park, August 30–September 1, 1971.

Fifty-fifth Annual Meeting, Las Vegas, Nevada, January 19–21, 1972.

The following is a list of the Sections of the Association with dates of future meetings so far as they have been reported to the Editorial Director.

ALLEGHENY MOUNTAIN
FLORIDA
ILLINOIS
INDIANA
IOWA
KANSAS
KENTUCKY
LOUISIANA-MISSISSIPPI
MARYLAND-DISTRICT OF COLUMBIA-VIRGINIA
METROPOLITAN NEW YORK
MICHIGAN
MISSOURI
NEBRASKA
NEW JERSEY
NORTH CENTRAL
NORTHEASTERN, Wellesley College, Wellesley, Massachusetts, November 27, 1971.
NORTHERN CALIFORNIA
OHIO
OKLAHOMA-ARKANSAS
PACIFIC NORTHWEST
PHILADELPHIA, Lafayette College, Easton, November 20, 1971.
ROCKY MOUNTAIN
SOUTHEASTERN
SOUTHERN CALIFORNIA
SOUTHWESTERN
TEXAS
UPPER NEW YORK STATE
WISCONSIN

FUTURE MEETINGS OF OTHER ORGANIZATIONS

AMERICAN ASSOCIATION FOR THE ADVANCEMENT OF SCIENCE, Philadelphia, December 26–31, 1971.

AMERICAN MATHEMATICAL SOCIETY, Pennsylvania State University, University Park, August 31–September 3, 1971.

AMERICAN SOCIETY FOR ENGINEERING EDUCATION

ASSOCIATION FOR COMPUTING MACHINERY, Chicago, August 3–5, 1971.

ASSOCIATION FOR SYMBOLIC LOGIC, Universidad Católica de Chile, Santiago, July 26–August 1, 1971.

CENTRAL ASSOCIATION OF SCIENCE AND MATHEMATICS TEACHERS, Detroit, Michigan, November 18–20, 1971.

FIBONACCI ASSOCIATION, College of the Holy Names, Oakland, California, November 13, 1971

INSTITUTE OF MATHEMATICAL STATISTICS, Fort Collins, Colorado, August 23–26, 1971.

MU ALPHA THETA, Pennsylvania State University, University Park, September 1, 1971.

NATIONAL COUNCIL OF TEACHERS OF MATHEMATICS, Chicago, Illinois, April 16–20, 1972.

OPERATIONS RESEARCH SOCIETY OF AMERICA, Disneyland Hotel, Los Angeles, October 27–29, 1971.

PI MU EPSILON, Pennsylvania State University, University Park, August 31–September 1, 1971.

SOCIETY FOR INDUSTRIAL AND APPLIED MATHEMATICS, University of Wisconsin, Madison, October 11–13, 1971.

GEORGE BANTA COMPANY, INC., MENASHA, WISCONSIN

THE AMERICAN MATHEMATICAL MONTHLY

(FOUNDED IN 1894 BY BENJAMIN F. FINKEL)

THE OFFICIAL JOURNAL OF

THE MATHEMATICAL ASSOCIATION OF AMERICA, INC.

VOLUME 78 NUMBER 7

CONTENTS

(Continued on inside cover)

AUGUST-SEPTEMBER 1971

NOTICE TO AUTHORS

Specialized research is usually unsuitable; see *Statement of Policy* (vol. 76, p. 2). Manuscript preparation: Please use the *Manual for Monthly Authors* (vol. 78, p. 1) and follow the format in current issues of the MONTHLY. Manuscripts should be typewritten, triple-spaced with wide margins; submit two copies and keep one for protection against loss. Each author receives 50 free reprints.

Backlog: Main Articles 7 months, Math. Notes 8 months, Research Problems 6 months, Classroom Notes 7 months, Math. Education 6 months.

EDITORIAL CORRESPONDENCE AND MAIN ARTICLES: to HARLEY FLANDERS, American Mathematical Monthly, Tel Aviv University, Ramat Aviv, Israel (see Notice, vol. 77, 1970, p. 555); NOTES, etc.: to the corresponding Associate Editor; ADVERTISING CORRESPONDENCE: to RAOUL HAILPERN, Mathematical Association of America, SUNY at Buffalo, Buffalo, N. Y. 14214; CHANGE OF ADDRESS and SUBSCRIPTIONS: to A. B. WILLCOX, Mathematical Association of America, 1225 Connecticut Ave., N.W., Washington, D. C. 20036.

Annual dues for members of the Association (including a subscription to the American Mathematical Monthly) are $12.50. For nonmembers the subscription price is $18.00.

PUBLISHED BY THE ASSOCIATION at Washington, D. C., and Menasha, Wisconsin during the months of January, February, March, April, May, June–July, August–September, October, November, December.

Second-class postage paid at Washington, D. C., and additional mailing offices. Acceptance for mailing at special rate of postage provided for in the Act of February 28, 1925, embodied in Paragraph 4, Section 538, P. L. and R., authorized April 1, 1926.

PRINTED IN THE UNITED STATES OF AMERICA

INTEGRAL INEQUALITIES INVOLVING A FUNCTION AND ITS DERIVATIVE

P. R. BEESACK, Carleton University

1. Introduction. In a recent paper on differential geometry, H. Flanders [**28**, pp. 712–13 and p. 715] wanted to show that continuously differentiable functions F_1, F_2 existed on $[0, 1]$ such that $F_i(0) = F_i(1) = 0$ for $i = 1, 2$, and

$$\int_0^1 x^2 F_1'^2(x)dx < 2\int_0^1 F_1^2(x)dx,$$

$$\int_0^1 x^6 F_2'^2(x)dx > 6\int_0^1 x^4 F_2^2(x)dx.$$

(The functions F_i were "variations" in the sense of the calculus of variations, and the above inequalities were needed in order to show that a corresponding second variation was negative or positive respectively.) The existence of F_1 is trivial since, as the author noted, one can let F_1 be the function whose graph is obtained from the polygonal line joining (0, 0) to $(\epsilon, 1)$ to (1, 0) by rounding off the corner, and taking ϵ sufficiently small. In a certain sense, the existence of F_2 is even more trivial! In fact it turns out that *any* nontrivial admissible F_2 (that is, $F_2 \in C^1[0, 1]$ with $F_2(0) = F_2(1) = 0$, but $F_2(x) \not\equiv 0$) satisfies the second of the above inequalities. More precisely, it was proved in [**28**] that for all admissible F_2,

$$\int_0^1 x^6 F_2'^2(x)dx \geqq \frac{25}{4}\int_0^1 x^4 F_2^2(x)dx. \tag{1.1}$$

We shall give Flanders' short elegant proof of (1.1) in Section 2, and analyze this proof to see, if possible, why it works. It will turn out that the same method of proof can be used to prove other integral inequalities similar to (1.1), namely those of the form

$$\int_a^b r\,|\,u'\,|^p dx \geqq \int_a^b s\,|\,u\,|^p dx \qquad (p > 1), \tag{1.2}$$

$$\int_a^b r\,|\,u'\,|^{p+1} dx \geqq \int_a^b s\,|\,u\,|^p\,|\,u'\,|\,dx \qquad (p > 0), \tag{1.3}$$

$$\int_a^b ru'^2 dx \geqq \int_a^b su^2 dx. \tag{1.4}$$

In these inequalities r and s will usually be positive continuous functions on the open interval (a, b), and the inequalities will hold for all $u \in C^1(a, b)$ which satisfy

Paul Beesack wrote his Ph.D. thesis at Washington University under Z. Nehari. He spent 5 years at McMaster before transferring to Carleton University in 1960. He spent a sabbatical leave at the Defence Research Board of Canada, and he has been active in preparing secondary school texts. His main research interest is differential and integral equations and inequalities. *Editor.*

certain other conditions. We call inequalities of the form (1.2) with boundary conditions $u(a)=0$ or $u(b)=0$ (or perhaps both, especially when p is an even integer) *Hardy type inequalities*. The prototype of such inequalities, having $a=0$, $b=\infty$, $r(x)\equiv x^{p-k}$, $s(x)\equiv(|k-1|/p)^p x^{-k}$, with $k\neq 1$, $p>1$ and $u(0)=0$ if $k>1$, or $u(x)=-\int_x^\infty u'dt$ if $k<1$, is called *Hardy's inequality* [**29**, Th. 330], although this name is sometimes reserved for the special case $k=p$ [**29**, Th. 327]. The inequality (1.1) is the special case $p=2$, $k=-4$, $u(x)=-\int_x^1 u'(t)dt$, with $u'(x)=0$ for $x>1$. A number of extensions and generalizations of Hardy's inequality and of (1.2) have been considered by Beesack [**7**], Boyd [**17**], Talenti [**53, 54**], and Tomaselli [**55**]. Another elementary proof of a theorem which includes Hardy's inequality has been given recently by Shum [**52**].

The inequality (1.3) is a generalization of *Opial's inequality* [**46**], which is the special case $a=0$, $p=1$, $r(x)\equiv 1$, $s(x)\equiv 4b^{-1}$, with $u(0)=u(b)=0$. A large number of papers have been written dealing with integral inequalities of the form (1.3) and its generalizations, of which we mention those of Olech [**45**], Beesack [**8**], Levinson [**39**], Mallows [**40**], Pederson [**47**], Holt [**30**], Yang [**58**], Redheffer [**50**], Calvert [**20**], Maroni [**41**], Boyd and Wong [**15**], Das and Beesack [**24**], Willett [**56**], Das [**25**], and Boyd [**16, 17**]. Of these, the papers [**39, 40, 45, 47**] are essentially simpler proofs of Opial's original result. Boyd's paper [**17**] is the most general of those cited, and also gives the best constants. Discrete analogues of Opial's inequality were considered by Wong [**57**] and Lee [**37**]; see also Beesack [**9**].

Inequalities of the form (1.4) involving boundary conditions on $u(x)$ at one or both of a and b, together with the orthogonality condition $\int_a^b sudx=0$, are said to be of the *Wirtinger type*. The case $a=-\pi$, $b=\pi$, $r=s\equiv 1$, with the conditions $u(-\pi)=u(\pi)$, $\int_{-\pi}^{\pi} udx=0$, is called *Wirtinger's inequality* in the standard references [**29**, p. 185] and [**1**, p. 177]. However, as noted by Mitrinovič and Vasič in [**42**] and [**43**, p. 141], it appears that this designation is unjustified in that this inequality (as well as the more general (1.4)) had been considered by a number of other authors at an earlier date. Other inequalities of this (and similar) type have been considered more recently by Schmidt [**51**], Bellman [**2, 3**], Pleijel [**49**], Block [**13**], Beesack [**4, 5**], Coles [**21, 22**], Levin and Stečkin [**38**], and Diaz and Metcalf [**26**]. Discrete analogues of (1.4) involving sums and finite differences have been considered by Fan, Taussky, and Todd [**27**], Block [**14**], and especially Pfeffer [**48**].

For further references to integral inequalities of the general kind considered here, including those involving derivatives of higher order, see the discussion and extensive bibliographies in Mitrinovič [**43**, §2.23].

Sections 3, 4, and 5 of this paper deal with integral inequalities of type (1.2) and (1.3) respectively, using the completely elementary approach suggested in Section 2. Other related and more powerful inequalities are also discussed without proof. The same elementary method is used in Sections 6, 7, and 8 to obtain a variety of integral inequalities of the Wirtinger type (1.4), and additional inequalities of this kind are discussed in the final Section 9.

2. The inequality (1.1). Let u be any admissible function for (1.1) so that $u(0)=u(1)=0$, and $u\in C^1[0,1]$. Letting $z(x)=x^{5/2}u(x)$, we have $z(0)=z(1)=0$, and

$$u'(x) = x^{-5/2}z'(x) - \frac{5}{2}x^{-7/2}z(x),$$

$$\begin{aligned}\int_0^1 x^6u'^2dx &= \int_0^1\left\{xz'^2 - 5zz' + \frac{25}{4}x^{-1}z^2\right\}dx \\ &= \int_0^1 xz'^2dx - \frac{5}{2}z^2\Big|_0^1 + \frac{25}{4}\int_0^1 x^4u^2dx \\ &\geqq \frac{25}{4}\int_0^1 x^4u^2dx,\end{aligned} \tag{2.1}$$

completing Flanders' proof of (1.1). Moreover, we see that equality can hold in (1.1) only if $z'=0$ above, that is, only if $u(x)=cx^{-5/2}$. However, the only admissible such u is the trivial one with $c=0$. This leaves open the question as to whether the constant 25/4 appearing on the right side of (1.1) is best possible. To show that it is best possible, we denote the best (largest) constant by K, and first approximate the (inadmissible) extremal function $u_0(x)\equiv x^{-5/2}$ by the (inadmissible) function v defined by

$$v(x) = \begin{cases}\epsilon^{-7/2}x, & 0\leqq x\leqq \epsilon,\\ x^{-5/2}, & \epsilon\leqq x\leqq\delta,\\ -\delta^{-5/2}(1-\delta)^{-1}(x-1), & \delta\leqq x\leqq 1.\end{cases}$$

(The graph of v is obtained from the graph of u_0 by replacing the parts of the latter graph with abscissae near 0 and 1 by straight line segments making $v(0)=v(1)=0$.) By direct computation,

$$\int_0^1 x^6v'^2dx = \frac{1}{7}+\frac{25}{4}\ln(\delta/\epsilon)+a_1(\delta),$$

$$\int_0^1 x^4v^2dx = \frac{1}{7}+\ln(\delta/\epsilon)+a_2(\delta),$$

where a_1, a_2 are positive. By smoothing the corners of the graph of v, we obtain an admissible function v_ϵ (having continuous first derivative on $[0, 1]$) with

$$\int_0^1 x^6v_\epsilon'^2dx < \frac{1}{7}+\frac{25}{4}\ln(\delta/\epsilon)+2a_1(\delta),$$

$$\int_0^1 x^4v_\epsilon^2dx > \ln(\delta/\epsilon),$$

for fixed δ, and all ϵ with $0<\epsilon<\delta(<1)$. Since, by assumption,

$$\int_0^1 x^6 v_\epsilon'^2 dx \geqq K \int_0^1 x^4 v_\epsilon^2 dx,$$

it follows that

$$K < \frac{\frac{1}{7} + \frac{25}{4}\ln(\delta/\epsilon) + 2a_1(\delta)}{\ln(\delta/\epsilon)} \qquad \text{for } 0 < \epsilon < \delta.$$

Letting $\epsilon \to 0+$, we see that the best possible constant K is less than or equal to 25/4. By what was proved in (2.1), it follows that $K=25/4$.

The result (2.1) actually contains more than we set out to prove. That is, it is clear that the inequality (2.1) holds for all functions u which are absolutely continuous on $[0, 1]$ with $u(1)=0$ and $\int_0^1 x^6 u'^2 dx < \infty$. For, all integrals appearing in (2.1) will still exist and be nonnegative, and $z^2(x)=x^5u^2(x)$ will still vanish at 0 and 1 for such u. If we had so enlarged our class of admissible functions, the proof that $K=25/4$ was the best possible constant would have been somewhat simpler. In the rest of this paper we shall, for simplicity, admit functions which are absolutely continuous and assume that the integrals are Lebesgue integrals. (One could interpret the integrals as absolutely convergent Riemann integrals and assume that the admissible functions u are such integrals of their derivatives.)

We now return to Flanders' proof of (2.1) and ask: why does the change of variable $u=yz$, with $y(x)=x^{-5/2}$, work? In order to see why the choice $y=x^{-5/2}$ produced the desired result, suppose we begin with the arbitrary change of variable $u=yz$, $u'=yz'+y'z$, and proceed formally as in (2.1). Then keeping an eye on (2.1), we have

$$\begin{aligned}\int_0^1 x^6u'^2dx &= \int_0^1 x^6y^2z'^2dx + 2\int_0^1 x^6yy'zz'dx + \int_0^1 x^6y'^2z^2dx \\ &\geqq x^6yy'z^2\Big|_0^1 - \int_0^1 z^2\{y(x^6y')' + x^6y'^2\}dx + \int_0^1 x^6y'^2z^2dx \\ &= x^6yy'z^2\Big|_0^1 - \int_0^1 yz^2(x^6y')'dx.\end{aligned}$$

Ignoring the boundary terms for now we see that, in order to obtain a result resembling (2.1), we want to choose y so that

$$-yz^2(x^6y')' = \frac{25}{4}x^4u^2.$$

Using $yz=u$, this means that y must be a solution of the second-order, linear, self-adjoint differential equation

$$(x^6y')' + \frac{25}{4}x^4y = 0. \tag{2.2}$$

This equation can be written as an Euler (or Cauchy) type equation, which can be reduced to a linear equation with constant coefficients by the change of variable $x=e^t$. In this way the general solution of (2.2) is found to be $y=x^{-5/2}(A+B\ln x)$. The boundary terms above become

$$x^6\left(\frac{y'}{y}\right)u^2 = x^5\frac{B-\frac{5}{2}A-\frac{5}{2}B\ln x}{A+B\ln x}u^2(x),$$

and become zero as $x\to 0+$ and at $x=1$, for *all* admissible functions u and all possible choices of A, B. In order to ensure that $z=y^{-1}u=x^{5/2}u(x)/(A+B\ln x)$ is continuous on (0, 1), the constants A and B must satisfy $AB\leqq 0$. Each of the above integrals will exist for any such choice of A, B. The choice $A=1$, $B=0$ is, of course, the simplest. However, the choice $A=0$, $B=-1$, leading to the (inadmissible) extremal function $u_1(x)=-x^{-5/2}\ln x$, would have simplified slightly the proof that the constant $K=25/4$ is best possible in (2.1), since now $u_1(1)=0$.

The differential equation (2.2) is also related to the inequality (2.1) when written in the form

$$\int_0^1 F(x, u, u')\,dx = \int_0^1\left\{x^6u'^2 - \frac{25}{4}x^4u^2\right\}dx \geqq 0, \tag{2.3}$$

by noting that it is the Euler-Lagrange differential equation of the variational problem $\int_0^1 F(x, u, u')dx=\text{minimum}$, that is,

$$\frac{d}{dx}F_{u'} - F_u = 2\left\{(x^6u')' + \frac{25}{4}x^4u\right\} = 0. \tag{2.4}$$

Nevertheless, as noted in [**29**, p. 174 ff.], the difficulties involved in a variational approach to inequalities of the type noted in Section 1 are considerable. In the following sections we will show how the above elementary method—consisting of setting $u=yz$, where y is an appropriate solution of the corresponding Euler-Lagrange equation—leads to relatively easy proofs of the inequalities in question. My thesis supervisor, Z. Nehari, first showed me this method in 1956, and mentioned that he thought it might date back to Jacobi! In this connection, Professor B. Schwarz of the Technion, Haifa, has recently pointed out that in a footnote on p. 52 of the 1904 edition (Dover reprint, 1961) of *Lectures on the Calculus of Variations*, O. Bolza also attributes this method to Jacobi in an article on the transformation of the second variation published in the *Journal für Mathematik*, Vol. XVII (1837), p. 68. In Bolza's book, the technique is used (pp. 53–54) to give a "second" proof of the positive character of the second variation, that is of a certain integral inequality arising in the classical calculus of variations.

3. Inequalities of Hardy type (1.2). If we assume that u' and u are positive, then the inequality (1.2) can be written in the form

$$\int_a^b \{ru'^p - su^p\}\,dx \geqq 0 \qquad (p > 1), \tag{3.1}$$

and the associated Euler-Lagrange differential equation (2.4) is

$$\frac{d}{dx}(ry'^{p-1}) + sy^{p-1} = 0 \qquad (p > 1). \tag{3.2}$$

For now we shall merely assume that r and s are positive and continuous on (a, b), and that (3.2) has a solution y which, together with y', is positive on (a, b). To admit functions u such that u' and u may change sign on (a, b), we note that if

$$u(x) = \int_a^x u'dt, \qquad \text{then } v(x) = \int_a^x |u'|\,dt \geqq |u(x)|,$$

with equality if and only if u' does not change sign. On the other hand, $v'(x) = |u'(x)|$, so that

$$\int_a^b s|u|^p dx \leqq \int_a^b sv^p dx \leqq \int_a^b rv'^p dx = \int_a^b r|u'|^p dx, \tag{3.3}$$

provided the second inequality holds. Moreover, the first inequality becomes an equality if and only if u' does not change sign on (a, b). (Precisely the same results hold if $u(x) = -\int_x^b u'dt$ and $v(x) = \int_x^b |u'|\,dt$, provided v' is replaced by $(-v')$ in (3.3).)

Now set $v = yz$, so $v' = yz' + zy'$. Then v, v', and z are also nonnegative on (a, b), along with y and y'. In place of the first step in (2.1) we proceed (hopefully) as follows, using (3.2):

$$\begin{aligned}
\int_a^b rv'^p dx &= \int_a^b r(yz' + zy')^p dx \geqq p\int_a^b r(yz')(zy')^{p-1}dx + \int_a^b r(zy')^p dx \\
&= ryy'^{p-1}z^p\Big|_a^b - \int_a^b z^p\{y(ry'^{p-1})' + ry'^p\}dx + \int_a^b r(zy')^p dx \\
&= ryy'^{p-1}z^p\Big|_a^b + \int_a^b sy^p z^p dx \\
&= r(y'/y)^{p-1}v^p\Big|_a^b + \int_a^b sv^p dx.
\end{aligned} \tag{3.4}$$

The validity of (3.4) depends, of course, on the existence of the integrals and the boundary terms, but above all on the equality

$$(a + b)^p \geqq pab^{p-1} + b^p \qquad \text{if } b \geqq 0,\ a + b \geqq 0,\ p > 1. \tag{3.5}$$

Obviously (3.5) is valid if $b=0$, and equality holds in this case if and only if $a=0$ as well. If $b>0$, then writing $x=a/b$ and dividing (3.5) through by b^p, we are to prove that $f(x)\equiv(1+x)^p-px-1\geqq 0$ for $x\geqq -1$, $p>1$. This follows from the fact that $f''(x)$ is nonnegative while $f'(x)=0$ only for $x=0$, so that $f(x)>f(0)=0$ for $x\geqq -1$, $x\neq 0$. *It follows that* (3.5) *is valid, and that equality holds if and only if* $a=0$. Moreover, since ry is positive on (a, b), it follows that *equality can hold in* (3.4) *only if* $z'\equiv 0$, *that is, only if* $v\equiv cy$ *for some constant* c.

A variety of boundary conditions could be imposed on u at a or b, depending on our assumptions (as yet unspecified) of the behaviour of the solution $y(x)$ of (3.2) as x tends to a or b. We note that the upper boundary term (at b) in (3.4) is nonnegative, so that for the validity of (3.3) it suffices to have only

$$\lim_{x\to a} r(x)[y'(x)/y(x)]^{p-1}v^p(x) = 0$$

for all v such that $v(x)=\int_a^x v'dt$ and $\int_a^b rv'^p dx<\infty$. We note that by [29, Th. 190], in order that a measurable function v' be integrable on $[a, X]$ whenever $\int_a^X rv'^p dx<\infty$, it is necessary that $r^{-1/p}\in L_q[a, X]$, that is, that $\int_a^X r^{-q/p}dx<\infty$, where $q^{-1}+p^{-1}=1$. (I overlooked this fact in [7] and [6], but it was pointed out by Tomaselli [55].) Hence, the assumption that $\int_a^X r^{-q/p}dx$ converge, although not essential in what follows, is a reasonable assumption.

THEOREM 3.1. *Suppose that* r *and* s *are positive and continuous on* (a, b), *where* $-\infty\leqq a<b\leqq\infty$, *and that the differential equation* (3.2) *has a solution* y *such that* $y'(x)>0$ *and* $y(x)=\int_a^x y'dt$ *for* $a<x<b$, *while*

$$r(x)[y'(x)/y(x)]^{p-1}\left(\int_a^x r^{-q/p}dt\right)^{p/q} = O(1) \qquad \text{as } x\to a^+, \tag{3.6}$$

where $p>1$ *and* $q=p/(p-1)$. *Let* u *be any function such that* $u(x)=\int_a^x u'dt$ *and* $\int_a^b r|u'|^p dx<\infty$. *Then*

$$\int_a^b s|u|^p dx \leqq \int_a^b r|u'|^p dx. \tag{3.7}$$

Moreover, equality holds in (3.7) *if and only if* $u\equiv cy$, *where* $c=0$ *unless* $\int_a^b ry'^p dx<\infty$ *and* $\lim_{x\to b-} r(x)y(x)y'^{p-1}(x)=0$.

Proof. As above, we set $v=\int_a^x |u'|dt$. We first show that $r(y'/y)^{p-1}v^p\to 0$ as $x\to a^+$. This follows from Hölder's inequality, which gives

$$v(x) = \int_a^x v'dt = \int_a^x r^{-1/p}r^{1/p}v'dt \leqq \left(\int_a^x r^{-q/p}dt\right)^{1/q}\left(\int_a rv'^p dt\right)^{1/p},$$

so that

$$0\leqq r(x)[y'(x)/y(x)]^{p-1}v^p(x)\leqq\left(\int_a^x rv'^p dt\right)r(x)[y'(x)/y(x)]^{p-1}\left(\int_a^x r^{-q/p}dt\right)^{p/q}.$$

The result now follows from the hypothesis (3.6). If we use this, (3.4) now gives

$$\int_a^b rv'^p dx \geqq \overline{\lim_{x\to b-}}\, r(x)[y'(x)/y(x)]^{p-1}v^p(x) + \int_a^b sv^p dx,$$

whence $\int_a^b rv'^p dx \geqq \int_a^b sv^p dx$, where equality can hold only if $v\equiv cy$, and $\lim_{x\to b} ryy'^{p-1}=0$ if $c\neq 0$. Combining this with (3.3) we obtain (3.7).

By the preceding remarks together with that noted following (3.3), equality can hold in (3.7) only if $u\equiv cy$ (in which case u' does not change sign) and $\lim_{x\to b} ryy'^{p-1}=0$ if $c\neq 0$. If $\int_a^b ry'^p dx<\infty$, then $u=cy$ is admissible, $v=|u|$, and the boundary term at a in (3.4) is zero, so that, for $c\neq 0$, equality holds in (3.7) if and only if $\lim_{x\to b} ryy'^{p-1}=0$. This completes the proof of the theorem.

The inequality (3.7) is certainly *sharp* (that is, the unit constant factor on the right side of (3.7) can not be decreased) if both $\int_a^b ry'^p dx<\infty$ and $\lim_{x\to b} ryy'^{p-1}=0$. If $\int_a^b ry'^p dx<\infty$ but $\lim_{x\to b} ryy'^{p-1}\neq 0$, then (3.7) is not sharp in general. This can be seen by taking $p=2$, $r(x)=s(x)\equiv 1$, $y(x)=\sin x$, with $0=a<b<\pi/2$. In this case the hypotheses of the theorem are satisfied but $\lim_{x\to b} ryy'^{p-1}\neq 0$; one easily verifies that the unit constant can be reduced to $4b^2/\pi^2$.

If $\int_a^b ry'^p dx=\infty$, *then* (3.7) *is sharp if either*

$$\int_a^X ry'^p dx = \infty \quad \textit{and} \quad \underline{\lim}_{x\to a}\, r(x)y(x)y'^{p-1}(x) < \infty, \tag{3.8}$$

or

$$\int_X^b ry'^p dx = \infty \quad \textit{and} \quad \underline{\lim}_{x\to b}\, r(x)y(x)y'^{p-1}(x) < \infty. \tag{3.9}$$

To prove this, it suffices to take

$$u'(x) = \begin{cases} 0, & a \leqq x \leqq a', \\ y'(x), & a' < x < b', \\ 0, & b' \leqq x \leqq b, \end{cases}$$

where a', b' will be fixed later. Then $u(x)=y(x)-y(a')>0$ for $a'\leqq x\leqq b'$, and

$$u^p(x) = y^p(x)\left\{1-\frac{y(a')}{y(x)}\right\}^p \geqq y^p(x)\left\{1-p\frac{y(a')}{y(x)}\right\},\quad a'\leqq x\leqq b'.$$

(This inequality is the special case of (3.5) obtained by taking $a=-y(a')/y(x)$ and $b=1$.) It follows that

$$\int_a^b su^p dx > \int_{a'}^{b'} su^p dx \geqq \int_{a'}^{b'} sy^p dx - py(a')\int_{a'}^{b'} sy^{p-1}dx.$$

On the other hand, by (3.2) we see that $\int_{a'}^{b'} sy^{p-1}dx = -ry'^{p-1}\big|_{a'}^{b'}$ and that

$$\int_{a'}^{b'} sy^p dx = -\int_{a'}^{b'} y(ry'^{p-1})' dx = -ryy'^{p-1}\Big|_{a'}^{b'} + \int_{a'}^{b'} ry'^p dx.$$

Hence,

$$\begin{aligned}\int_a^b su^p dx &> \int_{a'}^{b'} ru'^p dx + py(a')ry'^{p-1}\Big|_{a'}^{b'} - ryy'^{p-1}\Big|_{a'}^{b'} \\ &> \int_{a'}^{b'} ru'^p dx - pr(a')y(a')y'^{p-1}(a') - r(b')y(b')y'^{p-1}(b') \\ &> (1-\delta)\int_{a'}^{b'} ru'^p dx = (1-\delta)\int_a^b ru'^p dx,\end{aligned}$$

provided that

$$pr(a')y(a')y'^{p-1}(a') + r(b')y(b')y'^{p-1}(b') < \delta \int_{a'}^{b'} ry'^p dx. \tag{3.10}$$

If now (3.8) holds, then given any $\delta > 0$, (3.10) is satisfied by taking b' arbitrarily and a' appropriately close to a. The same argument applies if (3.9) holds, proving that (3.7) is sharp.

In order to obtain both parts of Hardy's inequality ($k<1$ as well as $k>1$) we require a theorem which is essentially the same as Theorem 3.1 but with the roles of a and b interchanged, and the Euler-Lagrange equation (3.2) replaced by

$$\frac{d}{dx}\{r(-y')^{p-1}\} - sy^{p-1} = 0. \tag{3.2'}$$

THEOREM 3.2. *Suppose that r and s are positive and continuous on (a, b), wherd $-\infty \leqq a < b \leqq \infty$, and that equation (3.2′) has a solution y such that $y'(x) < 0$ ane $y(x) = -\int_x^b y'dt$ for $a < x < b$, while*

$$r(x)[-y'(x)/y(x)]^{p-1}\left(\int_x^b r^{-q/p}dt\right)^{p/q} = O(1) \quad as \quad x \to b-, \tag{3.6'}$$

where $p>1$, $q = p/(p-1)$. Let u be any function such that $u(x) = -\int_x^b u'dt$ and $\int_a^b r|u'|^p dx < \infty$. Then

$$\int_a^b s\,|u|^p dx \leqq \int_a^b r\,|u'|^p dx. \tag{3.7'}$$

Equality holds in (3.7′) if and only if $u \equiv cy$, where $c=0$ unless both $\int_a^b r(-y')^p dx < \infty$ and $\lim_{x\to b} r(x)y(x)[-y'(x)]^{p-1} = 0$. The inequality (3.7′) is sharp if either

$$\int_a^X r(-y')^p dx = \infty \quad and \quad \varliminf_{x\to a} r(x)y(x)[-y'(x)]^{p-1} < \infty, \tag{3.8'}$$

or

$$(3.9') \qquad \int_X^b r(-y')^p dx = \infty \quad \textit{and} \quad \varliminf_{x\to b} r(x)y(x)[-y'(x)]^{p-1} < \infty.$$

The proof of this theorem is the same as that of Theorem 3.1 except that we now set $v(x) = \int_x^b |u'| dt$ (so v' is replaced by $(-v')$ as noted earlier), and (3.4) becomes

$$(3.4') \qquad \int_a^b r(-v')^p dx \geqq - r(-y'/y)^{p-1}v^p \Big|_a^b + \int_a^b sv^p dx.$$

Hardy's inequality is the special case of Theorem 3.1 (for $k>1$) and Theorem 3.2 (for $k<1$) obtained by taking $a=0$, $b=\infty$, $r(x) \equiv x^{p-k}$, $s(x) \equiv (|k-1|/p)^p x^{-k}$, $y(x) = x^{(k-1)/p}$. Here,

$$r^{-q/p}(x) = x^{-(p-k)/(p-1)}, \qquad \int_0^x r^{-q/p} dt = c_1 x^{(k-1)/(p-1)},$$

$$r(y'/y)^{p-1}\left(\int_0^x r^{-q/p} dt\right)^{p/q} = \left(\frac{k-1}{p}\right)^{p-1},$$

or

$$\int_x^\infty r^{-q/p} dt = c_2 x^{(k-1)/(p-1)}, \qquad r(-y'/y)^{p-1}\left(\int_x^\infty r^{-q/p} dt\right)^{p/q} = \left(\frac{1-k}{p}\right)^p,$$

according as $k>1$ or $k<1$. The hypotheses of Theorems 3.1 and 3.2 are thus satisfied and, since $\int_0^\infty r|y'|^p dx = \infty$, equality holds in (3.7) or (3.7′) only if $u \equiv 0$. On the other hand, since

$$r(x)y(x)\,|\,y'(x)\,|^{p-1} \equiv (\,|\,k-1\,|\,/p)^{p-1},$$

it follows that Hardy's inequality is sharp.

Other special cases of the inequalities (3.7) or (3.7′) are given in [**29**, Th. 256] and [**7**, p. 51]. In the latter paper integral inequalities of this type with $p<0$, and with $0<p<1$ (when the direction of the inequality sign is reversed), are also considered, as well as a discussion of the case that $p=2n$ (when the equations (3.2) and (3.2′) coincide). An interesting special case of the latter kind is the inequality

$$(3.11) \qquad \int_{-1}^1 \frac{u^2 dx}{(1-x^2)^2} \leqq \int_{-1}^1 u'^2 dx \quad \textit{if} \quad u(\pm 1) = 0,$$

with strict inequality unless $u \equiv 0$. This inequality is due to Nehari [**44**], and has applications to the theory of complex linear second order differential equations, and to the theory of univalent functions. The case $p=2$ of (1.2) with the boundary conditions $u(a) = u(b) = 0$ also leads to an elementary proof of the

minimum property of the least positive eigenvalue of the Sturm-Liouville problem

$$(ry')' + (q + \lambda s)y = 0, \qquad y(a) = y(b) = 0. \tag{3.12}$$

This is done for the case $q \equiv 0$, $r \equiv 1$ by Benson [**11**, p. 296] by a different elementary method which leads to a large number of other useful inequalities. The general case of (3.12) was dealt with in [**6**] and included elementary proofs of the variational characterization of the nth positive eigenvalue of (3.12), and of the Courant [**23**, p. 463] maximum-minimum properties of these eigenvalues.

Following some special cases considered by Talenti [**53, 54**], Tomaselli [**55**] dealt with the inequality (3.7) from a completely different point of view to that used here, or that used by me in [**7**], where an auxiliary Riccati-like equation was used rather than the substitution $u = yz$. Assuming only that r and s are nonnegative measurable functions on $(0, a)$ such that $\int_0^x r^{-q/p}dt$ exists for $0 \leqq x < a$, let C denote the best possible constant in order that the inequality

$$\int_0^a su^p dx \leqq C \int_0^a ru'^p dx \qquad (p > 1) \tag{3.13}$$

be valid for all functions $u(x) = \int_0^x u'dt$ with $u' \geqq 0$ on $(0, a)$. Tomaselli obtained the following upper bound for C, namely

$$C \leqq K = \frac{1}{p-1} \inf_{f \in F} \sup_{0<x<a} \frac{\int_0^x s(t) \left\{ f(t) + \int_0^t r^{-q/p} ds \right\}^p dt}{f(x)}, \tag{3.14}$$

where F is the class of all measurable strictly positive functions on $(0, a)$, and showed that $C = K$ if equality is attained in (3.13) for some nontrivial function u. By using special choices of u and f, Tomaselli also obtained a number of excellent lower and upper bounds for C.

The most general results obtained so far which include inequalities of Hardy type were given by Boyd in [**17**]. His results also include generalizations of the Opial type inequalities (1.3) and will be stated in Section 5.

Hardy's inequality has been applied to a variety of problems in recent years. For example, R. A. Hunt and G. Weiss [**33**] used the inequality to give a short, elegant proof of Marcinkiewicz's theorem on the interpolation of quasi-linear operators acting on L_p-spaces, and Hunt [**32**] used the same method to extend Marcinkiewicz's interpolation theorem to spaces of Lorentz type. In a very recent paper [**18**], D. Boyd used Hardy's inequality to prove a result on the osculatory packing of spheres in Euclidean n-dimensional space.

4. Inequalities of Opial type (1.3). If we again assume that u' and u are both positive, the inequality (1.3) becomes

$$\int_a^b \{ru'^{p+1} - su^p u'\} dx \geqq 0 \qquad (p > 0), \tag{4.1}$$

and the corresponding Euler-Lagrange equation (2.4) becomes

$$\frac{d}{dx}\{(p+1)ru'^p - su^p\} + psu^{p-1}u' = 0,$$

or

$$(ry'^p)' = (p+1)^{-1}s'y^p. \tag{4.2}$$

Here we shall assume that r and s are positive with r and s' continuous on (a, b), and that (4.2) has a solution y such that y and y' are positive on (a, b). (This is, of course, a strong assumption. In [15] Boyd and Wong essentially make this same assumption and obtain the following results with (a, b) replaced by the compact $[a, b]$. In [16, p. 386] Boyd introduces a factor λ on the right side of (4.2) and proves the existence of a solution of the resulting eigenvalue problem with appropriate boundary conditions. In this way he obtains best possible constants in inequalities of Opial type without making strong *ad hoc* assumptions such as we make here.) As in Section 3, if $u(x) = \int_a^x u'dt$ and $v(x) = \int_a^x |u'|dt$, then

$$\int_a^b s|u|^p|u'|dx \leqq \int_a^b sv^pv'dx (\leqq) \int_a^b rv'^{p+1}dx = \int_a^b r|u'|^{p+1}dx, \tag{4.3}$$

provided the second inequality holds. Again, setting $v = yz$, $v' = yz' + zy'$, we have on using (4.2),

$$\begin{aligned}\int_a^b rv'^{p+1}dx &= \int_a^b r(yz' + zy')^{p+1}dx \\ &\geqq (p+1)\int_a^b ryy'^pz^pz'dx + \int_a^b r(y'z)^{p+1}dx \\ &= ryy'^pz^{p+1}\Big|_a^b - (p+1)^{-1}\int_a^b s'(yz)^{p+1}dx,\end{aligned}$$

or

$$\int_a^b rv'^{p+1}dx \geqq \{r(y'/y) - (p+1)^{-1}s\}v^{p+1}\Big|_a^b + \int_a^b sv^p\, v'dx. \tag{4.4}$$

Assuming the existence of the integral on the left side and of the boundary terms, (4.4) is valid with equality holding if and only if $z' \equiv 0$, or $v \equiv cy$ for some constant c. Once again we use Tomaselli's remark that $\int_a^X r^{-Q/P}dt = \int_a^X r^{-1/p}dt$ must exist (here, $P = p+1 > 1$, and $Q = P/(P-1) = (p+1)/p$) if v' is to be integrable on $[a, X]$ for every v such that $\int_a^X rv'^{p+1}dx < \infty$. This suggests the boundary conditions of the following theorem:

THEOREM 4.1. *Suppose that r and s are positive with r and s' continuous on (a, b), where $-\infty \leqq a < b \leqq \infty$, and that equation* (4.2) *has a solution y such that*

$y'(x)>0$ and $y(x)=\int_a^x y'dt$ for $a<x<b$, while

$$r(x)[y'(x)/y(x)]^p \geqq (p+1)^{-1}s(x) \quad \textit{for } x \textit{ near } b \quad (x>x_0), \textit{ and} \tag{4.5}$$

$$\{r(x)[y'(x)/y(x)]^p-(p+1)^{-1}s(x)\}\left(\int_a^x r^{-1/p}dt\right)^p = O(1) \textit{ as } x\to a+. \tag{4.6}$$

Let u be any function such that $u(x)=\int_a^x u'dt$ and $\int_a^b r|u'|^{p+1}dx<\infty$. Then

$$\int_a^b s\,|u^pu'|\,dx \leqq \int_a^b r\,|u'|^{p+1}dx \quad (p>0). \tag{4.7}$$

Equality holds in (4.7) if and only if $u\equiv cy$, where $c=0$ unless both $\int_a^b ry'^{p+1}dx<\infty$ and $\lim_{x\to b} y(x)\{r(x)y'^p(x)-(p+1)^{-1}s(x)y^p(x)\}=0$. The inequality (4.7) is sharp if either

$$\int_a^X ry'^{p+1}dx=\infty \quad \textit{and} \quad \varliminf_{x\to a} ryy'^p<\infty, \tag{4.8}$$

or

$$\int_X^b ry'^{p+1}dx=\infty,\quad \varliminf_{x\to b} y\{ry'^p-(p+1)^{-1}sy^p\}<\infty, \textit{ and } ry'^p=O(1) \tag{4.9}$$

as $x\to b$.

Proof. The proof is much the same as that of Theorem 3.1 so we shall give only an outline. The inequality (4.7) follows from (4.3) and (4.4), together with the fact that (4.5) implies that the upper boundary term (at b) in (4.4) is nonnegative, while the term at a vanishes by (4.6) and

$$v^{p+1}(x) \leqq \left(\int_a^x rv'^{p+1}dt\right)\left(\int_a^x r^{-1/p}dt\right)^p. \tag{4.10}$$

Equality can hold in (4.7) only if it holds at both places in (4.3), hence only if $u=cy$, $v=|c|y$, and the boundary terms in (4.4) vanish. If y is admissible ($\int_a^b ry'^{p+1}dx<\infty$) then, as just noted, the boundary term at a vanishes for $v=|c|y$; but if $c\neq 0$, the boundary term at b vanishes if and only if

$$\lim_{x\to b} y\{ry'^p-(p+1)^{-1}sy^p\}=0.$$

To prove the sharpness of (4.7) we proceed as in the proof of Theorem 3.1, defining the admissible function u precisely as before, so that

$$u(x)=y(x)\{1-[y(a')/y(x)]\} \text{ for } a'\leqq x\leqq b'.$$

If $p\geqq 1$, we may proceed as before. If $0<p<1$, then, since $(1-s)^p\geqq 1-s$ for $0\leqq s\leqq 1$, we have $u^p(x)\geqq y^p(x)-y(a')y^{p-1}(x)$. Hence for all $p>0$,

$$u^p(x)\geqq y^p(x)-\alpha y(a')y^{p-1}(x) \quad \text{for } a'\leqq x\leqq b',$$

where $\alpha=\max(p, 1)$. Using (4.2) we obtain

$$(4.11)\qquad -\int_{a'}^{b'} sy^{p-1}y'dx = \frac{p+1}{p}\left\{ry'^p - (p+1)^{-1}sy^p\right\}\Big|_{a'}^{b'}$$

and

$$\int_{a'}^{b'} sy^py'dx = \int_{a'}^{b'} ry'^{p+1}dx - y\left\{ry'^p - (p+1)^{-1}sy^p\right\}\Big|_{a'}^{b'}.$$

Thus, since s and $u'(=y')$ are positive on (a', b'), on setting $g=ry'^p-(p+1)^{-1}sy^p$, we have

$$\begin{aligned}\int_{a'}^{b'} su^pu'dx &\geqq \int_{a'}^{b'} sy^py'dx - \alpha y(a')\int_{a'}^{b'} sy^{p-1}y'dx\\ &= \int_{a'}^{b'} ry'^{p+1}dx - y(x)g(x)\Big|_{a'}^{b'} + \alpha y(a')\frac{p+1}{p}g(x)\Big|_{a'}^{b'}\\ &\geqq (1-\delta)\int_{a'}^{b'} ry'^{p+1}dx = (1-\delta)\int_a^b ru'^{p+1}dx,\end{aligned}$$

or $\int_a^b su^pu'dx>(1-\delta)\int_a^b ru'^{p+1}dx$ $(0<\delta<1)$, provided a', b' are chosen so that

$$(4.12)\qquad y(x)g(x)\Big|_{a'}^{b'} - \alpha y(a')\frac{p+1}{p}g(x)\Big|_{a'}^{b'} \leqq \delta\int_{a'}^{b'} ry'^{p+1}dx.$$

Unfortunately, as can be seen from (4.11), g is a decreasing function. By regrouping the terms in (4.12) and using (4.5), we see that it suffices to choose a', $b'(>x_0)$ such that

$$y(b')g(b') + \left(\alpha\frac{p+1}{p} - 1\right)r(a')y(a')y'^p(a') + \alpha(p+1)y(a')r(b')y'^p(b') \leqq \delta\int_{a'}^{b'} ry'^{p+1}dx.$$

By (4.8) or (4.9) this may always be done, and the proof of sharpness is complete.

REMARK. In case g becomes (and hence stays) negative for x near b, the hypothesis (4.5) can be replaced by the two assumptions

$$(4.13)\qquad \int_a^b r^{-1/p}dt < \infty \quad\text{and}\quad \lim_{x\to b}\left\{r(y'/y)^p - (p+1)^{-1}s\right\} = 0.$$

For, the first of (4.13), together with (4.10), shows that v is bounded on (a, b) if it is admissible. This fact, together with the second of (4.13), ensures that the upper boundary term in (4.4) has the value zero (and that the integral $\int_a^b sv^pv'dx$ exists). In this case, of course, equality holds in (4.7) for $u=cy$ when-

ever y is admissible. By using the fact that now $g(b')<0$, and rearranging the terms in (4.12), one can show that when y is not admissible, (4.7) is now sharp if either (4.8) holds or

$$\int_X^b ry'^{p+1}dx = \infty, \quad \text{but} \quad \varliminf_{x\to b} (sy^p - ry'^p) < \infty. \tag{4.14}$$

Perhaps the most interesting special case of (4.7) is the inequality

$$\int_0^b |u^p u'|\, dx \leqq \frac{b^p}{p+1}\int_0^b |u'|^{p+1}dx \qquad (p>0), \tag{4.15}$$

valid for any u which is absolutely continuous on $[0, b]$ *with* $u(0)=0$, *equality holding if and only if* $u(x)\equiv cx$. It may be of interest to note that for (4.15) *both* of the hypotheses (4.5) and (4.13) are satisfied. The case $p=1$ of (4.15) is the original inequality of Opial [**46**]. The first proof of (4.15) appears to be that of Yang [**58**] who proved that *if* $p\geqq 0$, $q\geqq 1$, *then*

$$\int_0^b |u|^p |u'|^q dx \leqq \frac{qb^p}{p+q}\int_0^b |u'|^{p+q}dx \tag{4.16}$$

for any u which is absolutely continuous on $[0, b]$ *with* $u(0)=0$. (Yang stated his result only for $p\geqq 1$, $q\geqq 1$ but his proof is valid for $p\geqq 0$, $q\geqq 1$. The result is sharp only for $q=1$.) An earlier paper by Hua [**31**] proved (4.15) but only in case p is a positive integer. The inequality (4.15) is included in a (later) generalization of Calvert [**20**], and about the same time a short direct proof of (4.15) was given by Wong [**57**]. In addition the result is contained in the cited paper [**15**] of Boyd and Wong who essentially proved a "compact" case of Theorem 4.1 by an elementary method using an associated Riccati-like equation. Other examples of Theorem 4.1 are also given in [**15**, (8) and (9)].

By taking $s\equiv 1$, replacing r by $(p+1)^{-1}(\int_a^b r^{-1/p}dx)^p r$, and taking $y=\int_a^x r^{-1/p}dt$ in Theorem 4.1, we obtain the inequality

$$\int_a^b |u^p u'|\, dx \leqq \frac{1}{p+1}\left(\int_a^b r^{-1/p}dx\right)^p \int_a^b r|u'|^{p+1}dx, \qquad (p>0), \tag{4.17}$$

valid for any u such that $u(x)=\int_a^x u'dt$, *with equality holding if and only if* $u\equiv cy$. Here a or b (or both) may be infinite provided $\int_a^b r^{-1/p}dt<\infty$. The case $p=1$ of (4.17) was proved in [**8**, Th. 1] under slightly more restrictive hypotheses on u, and specific examples are given there; a (different) generalization of this was given by Maroni [**41**]. The inequality (4.17) was essentially obtained by Das and Beesack [**24**, (18)].

5. Other generalizations of Opial's inequality. The elementary method of proof (4.4) used in proving Theorem 4.1 may also be used to prove a more general inequality similar to Yang's result (4.16), involving two parameters p, q, but for $0<q\leqq 1$, $p+q\geqq 1$. Although we shall not give details or state any the-

orem, we shall indicate the principal step corresponding to (4.4); the interested reader may find and prove the corresponding theorem. In place of the differential equation (4.2) we consider

$$\frac{d}{dx}\{(p+q)ry'^{p+q-1} - qsy^py'^{q-1}\} + psy^{p-1}y'^q = 0, \tag{5.1}$$

and assume this equation has a solution $y(x)=\int_a^x y'dt$, with $y'(x)>0$ on (a, b). Using $p+q\geqq 1$ and proceeding as in (4.4), we first obtain

$$\begin{aligned}\int_a^b rv'^{p+q}dx &\geqq ryy'^{p+q-1}z^{p+q}\Big|_a^b - \int_a^b z^{p+q}y(ry'^{p+q-1})'dx \\ &= \left(ryy'^{p+q-1} - \frac{q}{p+q}\,sy^{p+1}y'^{q-1}\right)z^{p+q}\Big|_a^b \\ &\quad + \int_a^b sv^p\{q(yz')(y'z)^{q-1} + (y'z)^q\}\,dx.\end{aligned}$$

We now use the fact that $0<q\leqq 1$, in which case one can verify that $(a+b)^q \leqq qab^{q-1}+b^q$ *if* $b\geqq 0$, $a+b\geqq 0$. Hence the preceding inequality gives

$$\int_a^b rv'^{p+q}dx \geqq \int_a^b sv^pv'^q dx + \left\{r(y'/y)^{p+q-1} - \frac{q}{p+q}\,s(y'/y)^{q-1}\right\}v^{p+q}\Big|_a^b. \tag{5.2}$$

Except for the fact that the above analysis does not require that a solution y of (5.1) be admissible (that is, satisfy $\int_a^b ry'^{p+q}dx<\infty$ and appropriate boundary conditions) any theorem so obtained will be included in the following theorem of D. W. Boyd [**17**].

THEOREM 5.1 (Boyd). *Suppose that r, $s\in C^1(a, b)$, that $s(x)>0$ a.e. and $r(x)>0$ for $a<x<b$, that $p>0$, $k\geqq 1$, $0\leqq q<k$, and that the operator T_1 defined by $T_1f(x) = s^{1/p}(x)r^{-1/p}(x)\int_a^x f(t)dt$ is compact from $L_r^k\to L_r^\alpha$, where $\alpha=pk/(k-q)$ and L_r^β is the set of measurable functions on (a, b) such that $\int_a^b r|f|^\beta dx<\infty$. Then the following eigenvalue problem* (P) *has solutions (y, λ) with $y\in C^2(a, b)$ and $y'(x)>0$, $y(x)>0$ in (a, b).*

$$(\mathrm{P})\begin{cases}(\mathrm{i})\ \dfrac{d}{dx}\{\lambda kry'^{k-1} - qsy^py'^{q-1}\} + psy^{p-1}y'^q = 0,\\[2ex] (\mathrm{ii})\ \lim\limits_{x\to a} y(x) = 0 \quad \textit{and} \quad \lim\limits_{x\to b}(\lambda kry'^{k-1} - qsy^py'^{q-1}) = 0,\\[2ex] (\mathrm{iii})\ \displaystyle\int_a^b r|y'|^k dx = 1.\end{cases}$$

There is a largest value λ such that (P) *has a solution, and if λ^* denotes this value, then for any u such that $u(x)=\int_a^x u'dt$ with $u\in L_r^k$,*

$$\int_a^b s\,|u|^p\,|u'|^q dx \leqq \frac{k\lambda^*}{p+q}\left\{\int_a^b r\,|u'|^k dx\right\}^{(p+q)/k}. \tag{5.3}$$

Moreover, equality holds if and only if $u = cy$ a.e., where y is a solution of (P) *corresponding to $\lambda = \lambda^*$.*

Note that (5.3) includes both integral inequalities of Opial type ($k = p+1$, $q = 1$), and of Hardy type ($k = p \geqq 1$, $q = 0$). The proof of this very general and powerful theorem by Boyd consists of a skillful blending of hard (classical) and soft (functional) analysis, and should serve to eliminate any pejorative implications from the use of either of the adjectives in question! We note that (P)(i) reduces to (5.1) by setting $k = p+q$; however, our proof of (5.2) required $0 < q \leqq 1$, $p+q \geqq 1$, whereas Boyd's result is valid for all p, q such that $p > 0$, $q \geqq 0$, $p+q \geqq 1$. In the same paper [**17**, Th. 2] Boyd explicitly obtained the best possible constants for (5.3) with $a = 0$, $b = 1$, $r(x) = s(x) \equiv 1$, $p > 0$, $k \geqq 1$, $0 \leqq q \leqq k$. This gives the sharp version of Yang's inequality (4.16) which had been only slightly improved by Das and Beesack [**24**, (17)] for $p > 0$, $q > 0$, $p+q \geqq 1$. In [**24**] inequalities of the form (5.3)—with $k = p+q$ always—were considered using essentially only Hölder's inequality, but for a broad range of values of $p+q$, including negative values. In some cases the direction of the inequality sign is reversed as, for example, if $p < 0$, $q \geqq 1$, $0 < p+q < 1$, or $p < 0$, $p+q \geqq 1$, or $p > 0$, $p+q < 0$. In most cases the inequalities obtained in [**24**] are not sharp. J. Calvert had also considered some reversed inequalities of this type including the following [**20**, p. 75]:

$$\int_a^b |u^p u'|\,dx \geqq \frac{1}{p+1}\left(\int_a^b r^{-1/p}dx\right)^p \int_a^b r\,|u'|^{p+1}dx \qquad (-1 < p < 0), \tag{5.4}$$

valid for any u which is an integral, $u(x) = \int_a^x u'dt$, with equality holding only if $u(x) = c\int_a^x r^{-1/p}dt$, (cf. (4.17)).

Turning now to results of a somewhat different character, we note first that on replacing (a, x) by (x, b) throughout, we can obtain a companion theorem to Theorem 4.1 in the same way that Theorem 3.2 was obtained. Although the interested reader is invited to state and prove such a theorem, we shall not do so here, but only note that the results corresponding to (4.2)–(4.4) are now

$$[r(-y')^p]' = (p+1)^{-1}s'y^p, \qquad (p > 0), \tag{5.2$'$}$$

$$\int_a^b s\,|u^p u'|\,dx \leqq \int_a^b s v^p(-v')dx \leqq \int_a^b r(-v')^{p+1}dx = \int_a^b r\,|u'|^{p+1}dx, \tag{5.3$'$}$$

$$\int_a^b r(-v')^{p+1}dx \geqq \left\{-r(-y'/y)^p + (p+1)^{-1}s\right\}v^{p+1}\Big|_a^b + \int_a^b s v^p(-v')dx, \tag{5.4$'$}$$

where $u(x) = -\int_x^b u'dt$ and $v(x) = \int_x^b |u'|\,dt$. Note that (5.2$'$) and (4.2) coincide in case p is an even integer (or more generally, if p is a rational of the form $2n/(2m+1)$), so that results of the kind mentioned in Section 3 could be ob-

tained. Instead, we want to point out how one may obtain sharp inequalities corresponding to boundary conditions of the form $u(a)=u(b)=0$ *for arbitrary* $p>0$. (The same technique could be applied to the Hardy type inequalities or to Boyd's inequalities (5.3), but this has apparently not been done.) The idea is to split the interval (a, b) into two intervals $(a, X]$ and $[X, b)$; suppose that the two differential eigenvalue problems

$$(ry_1'^p)' = \lambda_1^{-1}(p+1)^{-1}s'y_1^p, \qquad [r(-y_2')^p]' = \lambda_2^{-1}(p+1)^{-1}s'y_2^p,$$

have solutions (y_1, λ_1) and (y_2, λ_2) on $(a, X]$ and $[X, b)$ respectively, with $y_1(a)=y_2(b)=0$, satisfying the appropriate boundary conditions of Theorem 4.1 (with b replaced by X and s by $\lambda_1^{-1}s$) and the corresponding conditions for y_2 on $[X, b)$. Writing $\lambda_1=\lambda_1(X)$, $\lambda_2=\lambda_2(X)$, one then obtains

$$\int_a^b s\,|\,u^pu'\,|\,dx \leqq \lambda_1(X)\int_a^X r\,|\,u'\,|^{p+1}dx + \lambda_2(X)\int_X^b r\,|\,u'\,|^{p+1}dx,$$

valid for admissible functions u such that $u(x)=\int_a^x u'dt=-\int_x^b u'dt$ (so $u(a)=u(b)=0$) and $\int_a^b r|u'|^{p+1}dx<\infty$. If, in addition, $X=X_0$ can be chosen on (a, b) such that $\lambda_1(X_0)=\lambda_2(X_0)$, then we obtain

$$\int_a^b s\,|\,u^pu'\,|\,dx \leqq \lambda_1(X_0)\int_a^b r\,|\,u'\,|^{p+1}dx, \qquad (p>0), \tag{5.5}$$

valid for all such admissible functions u. Inequalities of this kind in the case $p=1$ were considered by Calvert [**20**], Beesack [**8**], Yang [**58**]; Opial's original inequality was actually of this kind. The paper [**24**] of Das and Beesack deals systematically with inequalities of the type (5.5), one such example being

$$\int_a^b |\,u^pu'\,|\,dx \leqq \frac{1}{p+1}\left(\int_a^{X_0} r^{-1/p}dt\right)^p \int_a^b r\,|\,u'\,|^{p+1}dx, \qquad (p>0), \tag{5.6}$$

where X_0 is the unique solution of the equation $\int_a^{X_0} r^{-1/p}dt=\int_{X_0}^b r^{-1/p}dt$. Equality holds in (5.6) if and only if $u=A\int_a^x r^{-1/p}dt$ for $a\leqq x\leqq X_0$, and $u=B\int_x^b r^{-1/p}dt$ for $X_0\leqq x\leqq b$. (Compare the above constant with that appearing in (4.17).) Other inequalities involving the boundary conditions $u(a)=u(b)=0$ and *two* parameters p, q, were also given in [**24**] and [**58**], but these were not usually sharp.

We conclude this section by pointing out that some attempts have been made to obtain inequalities of the Opial and Hardy type (and their generalizations) involving higher-order derivatives. Among these we mention the papers of Janet [**34, 35**], Pfeffer [**48**], and Kim [**36**], who obtained inequalities of Hardy type (with $p=2$). For example, Kim proved the following sharp inequality: *if* $u\in C^n[a, b]$ *with* $u^{(i)}(a)=u^{(i)}(b)=0$ *for* $i=0, \cdots, n-1$, *then*

$$\int_a^b \{u^{(n)}(x)\}^2dx \geqq [(b-a)/2]^{2n}\left\{\prod_{k=0}^{n-1}(2k+1)^2\right\}\int_a^b \frac{u^2(x)dx}{(x-a)^{2n}(b-x)^{2n}}, \tag{5.7}$$

with equality if and only if $u\equiv 0$. The inequality (3.11) of Nehari is the special case $n=1$, $a=-1$, $b=1$ of (5.7), and the inequality in [4, (2.15)] is the corresponding special case with $n=2$. Kim needed the inequality (5.7) to obtain disconjugacy criteria for the complex differential equation $y^{(n)}+py=0$. D. Willett [56] proved that *if* $u\in C^{n-1}[a,b]$, *and* $u^{(n-1)}$ *is absolutely continuous with* $u^{(i)}(a)=0$ *for* $i=0,\cdots,n-1$, *then*

$$\int_a^b |u^{(n)}u|\,dx \leqq c_n(b-a)^n\int_a^b |u^{(n)}|^2dx, \tag{5.8}$$

for a constant $c_n\leqq 1/2$. Willett used (5.8) to give an elementary proof of the existence-uniqueness theorem for the general linear differential equation of order n. Das [25] obtained a better upper bound for the constant c_n in (5.8), and Boyd [16] explicitly determined the best possible value of c_n, showing incidentally that $c_n=\frac{1}{2}(b_n/n!)$, where $b_n\rightarrow\frac{1}{2}$ as $n\rightarrow\infty$. The most general results were indicated by Boyd in [17], who pointed out that his method leads to best possible constants in inequalities of the form (5.3) but with u' replaced by $u^{(n)}$, where $u^{(i)}(a)=0$ for $i=0,\cdots,n-1$, and $u^{(n-1)}$ is absolutely continuous on $[a,b]$. In [17, p. 382] Boyd also indicates that his method is applicable to inequalities involving two-point boundary conditions (and derivatives of higher order), but no details are given for any of the higher-order cases.

It is apparent from some of the examples cited that inequalities of the kind we have been considering have applications to differential equations in the real and complex domain. The long paper of J. M. Holt [30] uses inequalities of this kind to obtain non-oscillation theorems. In a recent Ph.D. thesis, J. Brink [19] systematically considered inequalities of the form

$$\|u\|_p \leqq K(n,p,q,Z)(b-a)^{n+(1/p)-(1/q)}\|u^{(n)}\|_q \qquad (1\leqq p,q\leqq\infty),$$

valid for all $u\in C^n[a,b]$ such that u has (at least) n zeros on $[a,b]$ at the points of certain specified sets Z. (Here, $\|f\|_p$ denotes the usual L_p norm of f. Some of the constants K are sharp, others are not.) These inequalities were then applied to obtain disconjugacy criteria for nth order differential equations.

6. Wirtinger type inequalities (I). In order to obtain integral inequalities of the form

$$\int_a^b ru'^2dx \geqq \int_a^b su^2dx \tag{6.1}$$

for functions u satisfying certain boundary conditions and an orthogonality condition $\int_a^b su\,dx=0$, we suppose that y is an "appropriate" nontrivial (that is, $y(x)\not\equiv 0$) solution of the associated Euler-Lagrange differential equation

$$(ry')'+sy=0. \tag{6.2}$$

As before, we also assume that r and s are positive and continuous on (a,b). In order to permit the orthogonality condition, we make the more general

change of variable

$$u - C = yz, \qquad u' = yz' + y'z, \tag{6.3}$$

(rather than $u = yz$) in the left side of (6.1) to obtain in the usual way the (formal) inequality

$$\int_a^b ru'^2 dx \geqq r(y'/y)(u - C)^2 \Big|_{a+}^{b-} + \int_a^b s(u - C)^2 dx, \tag{6.4}$$

with equality holding if and only if $z' \equiv 0$, or $u \equiv C + ky$. The formal steps leading to (6.4), as well as the result itself, are clearly valid for any constant C, and any function u which is locally absolutely continuous on (a, b) and has $\int_a^b ru'^2 dx < \infty$, *provided the solution y of* (6.2) *is not zero on* (a, b), *and the limits*

$$\begin{aligned} A &= \underline{\lim}_{x\to a+} r(x)[y'(x)/y(x)]\{u(x) - C\}^2, \\ B &= \overline{\lim}_{x\to b-} r(x)[y'(x)/y(x)]\{u(x) - c\}^2, \end{aligned} \tag{6.5}$$

are both finite. Moreover, (6.4) *is also valid if the solution y has a zero at a point $\bar{x} \in (a, b)$ and the limits* (6.5) *are finite, if u is locally absolutely continuous on (a, b) and, in addition,* $C = u(\bar{x})$. This follows on noting that by Schwarz's inequality we have

$$\{u(x) - C\}^2 = \left(\int_{\bar{x}}^x u' dt\right)^2 \leqq |x - \bar{x}| \left| \int_{\bar{x}}^x u'^2 dt \right|.$$

(Note that the integral on the right side exists since r is strictly positive and continuous on (a, b) and, by assumption, $\int_a^b ru'^2 dx < \infty$.) Hence, since

$$\lim_{x\to\bar{x}} \frac{r(x)y'(x)}{y(x)}(x - \bar{x}) = \lim_{x\to\bar{x}} \frac{r(x)y'(x)}{[y(x) - y(\bar{x})]/(x - \bar{x})} = r(\bar{x})$$

exists, it follows that $r(y'/y)(u - C)^2 \to 0$ as $x \to \bar{x}$, if $C = u(\bar{x})$.

We now expand the right side of (6.4) to obtain the inequality

$$\int_a^b ru'^2 dx \geqq B - A + \int_a^b su^2 dx - 2C \int_a^b su\, dx + C^2 \int_a^b sdx, \tag{6.6}$$

which is valid under the additional assumption that at least two (hence all three) of the integrals on the right side exist. (If $C = 0$, of course, no additional hypotheses are required.) Whenever either of the integrals $\int_a^b sdx$ or $\int_a^b su\, dx$ occurs in the statement of a theorem, the existence of the integral is assumed as part of the hypothesis. In case $a = -\infty$, we shall write $u(a)$ to denote $\lim_{x\to-\infty} u(x)$ and assume this limit exists (finite). A statement such as "*u is absolutely continuous on $[a, b)$*," both for finite a and for $a = -\infty$, is to be interpreted as meaning that u is an integral on $[a, b)$, that is, that

$$u(x) = u(a) + \int_a^x u'dt \quad \text{for } x \in (a, b),$$

with similar statements for $x=b$. Similarly, we shall always assume that solutions y of (6.2) are absolutely continuous on $[a, b)$ if the boundary value $y(a)$ appears in the statement of the theorem, with similar remarks for $y(b)$.

THEOREM 6.1. *Suppose that r and s are positive and continuous on (a, b), where $-\infty \leqq a < b \leqq \infty$, and that equation* (6.2) *has a solution y_1 with $y_1(a)=0$ and $y_1'(x)>0$ on (a, b), and*

$$r(x)[y_1'(x)/y_1(x)] \int_a^x r^{-1}dt = O(1) \quad \text{as } x \to a+. \tag{6.7}$$

Let u be absolutely continuous on $[a, b)$ with $\int_a^b ru'^2dt < \infty$ and

$$u(a) \int_a^b su\, dx \leqq 0. \tag{6.8}$$

Then

$$\int_a^b ru'^2dx \geqq \int_a^b su^2dx + u^2(a) \int_a^b sdx. \tag{6.9}$$

If either $\int_a^b ry_1'^2dx = \infty$, or $\alpha = \lim_{x\to a+} r(x)y_1'(x) = \infty$, or $\beta_1 = \lim_{x\to b-} r(x)y_1(x)y_1'(x) \neq 0$, equality holds in (6.9) *only for $u\equiv 0$. Otherwise equality holds in* (6.9) *if and only if either $u \equiv ky_1$, or*

$$u(x) = u(a) \left\{1 - \alpha^{-1}\left(\int_a^b sdx\right) y_1(x)\right\}, \tag{6.10}$$

where k is an arbitrary constant. (If we set $\beta = \lim_{x\to b-} r(x)y_1'(x)$, then the limits α, β always exist, with $0 \leqq \beta < \alpha \leqq \infty$; if $\beta > 0$, then $\beta_1 \neq 0$.)

Proof. The existence and properties of the limits α, β follow from the fact that $(ry_1')' = -sy_1$ is negative on (a, b), so the positive function ry_1' is strictly decreasing on (a, b). Also, $0 < \lim_{x\to b-} y(x) \leqq \infty$ holds, so that $\beta_1 > 0$ (possibly $\beta_1 = \infty$) if $\beta > 0$.

To prove (6.9) set $y=y_1$ in (6.4) and take $C=u(a)$. By Schwarz's inequality, we then have

$$\{u(x) - C\}^2 = \left(\int_a^x u'dt\right)^2 \leqq \left(\int_a^x r^{-1}dt\right)\left(\int_a^x ru'^2dt\right).$$

Using the first of (6.5) with $y=y_1$, and (6.7), we see that $A=0$. Moreover, $B\geqq 0$ since $r(y_1'/y_1)$ is positive on (a, b). The fact that B is finite follows from the fact that the two nonzero terms on the right side of (6.4) are *both* positive. It now follows from (6.6) that

$$\int_a^b ru'^2dx \geqq \int_a^b su^2dx + u^2(a)\int_a^b sdx - 2u(a)\int_a^b su\,dx + B,$$

so that (6.9) follows from (6.8).

Equality can hold in (6.9) only if $u(x)=u(a)+ky_1(x)$, $B=0$, and equality holds in (6.8). If $\int_a^b ry_1'^2dx=\infty$, then u is not admissible unless $k=0$ and, in this case, equality in (6.8) would imply that $u(a)=0$, so $u(x)\equiv 0$. Also, in order that $u(x)=u(a)+ky_1(x)$ be admissible, the integral in (6.8) must exist. But

$$\int_a^b s[u(a)+ky_1]dx = u(a)\int_a^b sdx - k\int_a^b (ry_1')'dx$$

$$= u(a)\int_a^b sdx - k[\beta - \lim_{x\to a+} r(x)y_1'(x)]$$

exists if and only if α is finite, unless $k=0$. Thus, as above, when $\alpha=\infty$, equality holds in (6.9) only for $u\equiv 0$. For $u(x)=u(a)+ky_1(x)$, one finds $B=k^2\ \overline{\lim}_{x\to b}\ r(x)y_1'(x)y_1(x)$, so if $\beta_1\neq 0$, then $B>0$ unless $k=0$, and we again see that equality holds in (6.9) only for $u\equiv 0$.

Finally, if α is finite, $\beta_1=0$, and $\int_a^b ry_1'^2dx<\infty$, then $\beta=0$, $B=0$, and for equality in (6.9) we require both $u(x)=u(a)+ky_1(x)$, and

$$[u(a)+ky_1(a)]\int_a^b s[u(a)+ky_1]dx = u(a)\left\{u(a)\int_a^b sdx + k\alpha\right\} = 0,$$

so that either $u(a)=0$ or $k=-\alpha^{-1}u(a)\int_a^b sdx$. If $u(x)\equiv ky_1(x)$, then u is clearly admissible, equality holds in (6.8) and hence also in (6.9). If u is given by (6.10), then as was just shown, u is also admissible with equality in (6.8) and (6.9).

Remark 1. In case $u(a)=0$ there is no need to require the existence of either of the integrals $\int_a^b s\,dx$ or $\int_a^b su\,dx$ as a (hidden) hypothesis. Moreover the reference to α becomes redundant, and in fact the theorem becomes a special case of Theorem 3.1.

Remark 2. By taking the same form of admissible function u as that used in Section 3, one can show, by precisely the same method, that if $\int_a^b ry_1'^2dx=\infty$, *the inequality* (6.9) *is sharp if either*

$$\int_a^X ry_1'^2dx = \infty \quad \textit{and} \quad \underline{\lim}_{x\to a+} r(x)y_1(x)y_1'(x) < \infty, \tag{6.11}$$

or

$$\int_X^b ry_1'^2dx = \infty \quad \textit{and} \quad \underline{\lim}_{x\to b-} r(x)y_1(x)y_1'(x) < \infty. \tag{6.12}$$

(Since $u(a)=0$, and $\int_a^b su\,dx$ exists for the function u in question, the sharpness in fact follows as a special case of that proved in Section 3.)

THEOREM 6.2. *Suppose that r and s are positive and continuous on (a, b), where $-\infty \leqq a < b \leqq \infty$, and that equation* (6.2) *has a solution y_2 with $y_2'(x) < 0$ on (a, b) and $y_2(\bar{x}) = 0$ at a (single) point $\bar{x} \in (a, b)$. Let u be locally absolutely continuous on (a, b) with $\int_a^b ru'^2dx < \infty$, and*

$$u(\bar{x}) \int_a^b su\,dx \leqq 0. \tag{6.13}$$

Then

$$\int_a^b ru'^2dx \geqq \int_a^b su^2dx + u^2(\bar{x}) \int_a^b s\,dx. \tag{6.14}$$

If $\int_a^b ry_2'^2dx = \infty$, or $\alpha_1 = \lim_{x\to a} r(x)y_2(x)y_2'(x) \neq 0$, or $\beta_1 = \lim_{x\to b} r(x)y_2(x)y_2'(x) \neq 0$, equality holds in (6.14) *only for $u \equiv 0$. Otherwise, equality holds in* (6.14) *if and only if $u \equiv ky_2$.*

Proof. First we note that ry_2' is negative and decreasing on $(a, \bar{x})$, and negative and increasing on $(\bar{x}, b)$, so that both the limits

$$\alpha = \lim_{x\to a+} r(x)y_2'(x), \qquad \beta = \lim_{x\to b-} r(x)y_2'(x),$$

exist (finite), and are nonpositive. *Clearly*, $\alpha_1 = 0 (\beta_1 = 0)$ *implies* $\alpha = 0 (\beta = 0)$.

To prove (6.14) we proceed as in the last theorem, but now taking $y = y_2$ and $C = u(\bar{x})$ in (6.4). Since ry_2'/y_2 is negative on $(a, \bar{x})$ and positive on $(\bar{x}, b)$, it follows that $-A \geqq 0$ and $B \geqq 0$. The inequality (6.14) now follows from (6.6) (with $C = u(\bar{x})$) and (6.13). Moreover, equality can hold in (6.14) only if $u(x) \equiv u(\bar{x}) + ky_2(x)$, $A = B = 0$, and equality holds in (6.13). The proof of the conditions for equality follows precisely as in the proof of Theorem 6.1, except that now the integral

$$\int_a^b s[u(\bar{x}) + ky_2]dx = u(\bar{x}) \int_a^b sdx - k(\beta - \alpha)$$

always exists, and reduces to $u(\bar{x})\int_a^b sdx$ if $\alpha_1 = \beta_1 = 0$.

REMARK 3. Again, if $u(\bar{x}) = 0$, the existence of the integrals $\int_a^b sdx$ and $\int_a^b su\,dx$ need not be assumed. By taking

$$u(x) = \begin{cases} y_2(a'), & a \leqq x \leqq a', \\ y_2(x), & a' \leqq x \leqq b', \\ y_2(b'), & b' \leqq x \leqq b, \end{cases}$$

one can show that *the inequality* (6.14) *is sharp if either* (6.12) *holds* (*with y_1 replaced by y_2*), *or if*

$$\int_a^X ry_2'^2\,dx = \infty \quad \textit{and} \quad \overline{\lim_{x\to a+}}\, r(x)y_2(x)y_2'(x) > -\infty. \tag{6.15}$$

The preceding two theorems are extensions of Theorems 1.1* and 1.2 of [4], and a number of examples (on compact intervals) are given there. The simplest, and probably the most useful of these, are the following inequalities:

(6.16) *If* $u' \in L_2$ *and* $u(0)\int_0^{\pi/2} udx \leqq 0$, *then*

$$\int_0^{\pi/2} u'^2dx \geqq \int_0^{\pi/2} u^2dx + (\pi/2)u^2(0),$$

with equality only for $u \equiv k \sin x$ *or* $u = u(0)(1 - (\pi/2)\sin x)$.

(6.17) *If* $u' \in L_2$ *and* $u(\pi/2)\int_0^{\pi} udx \leqq 0$, *then*

$$\int_0^{\pi} u'^2dx \geqq \int_0^{\pi} u^2dx + \pi u^2(\pi/2),$$

with equality only for $u \equiv k \cos x$.

When $u(0)=0$, the inequality (6.16) reduces to that in [**29**, p. 184]. The elementary proofs given in both [**4**] and [**29**] depend on an integral identity (related to a Riccati equation) rather than the change of variable (6.3), although the methods are essentially the same.

7. Wirtinger type inequalities (II). In the preceding section we dealt with those problems for which the extremal function, y, had only a single zero on $[a, b]$. In the present section we shall deal with problems where the extremal function has two zeros. From now on it will usually be more convenient to replace the interval (a, b) by an interval $(-a, a)$, symmetric about the origin.

THEOREM 7.1. *Suppose that* r *and* s *are positive and continuous on* $(-a, a)$, *where* $0<a\leqq\infty$, *and that equation* (6.2) *has a solution* y_1 *which is positive on* $(-a, a)$, *with* $y_1(\pm a)=0$, *and*

$$r(x)[y_1'(x)/y_1(x)]\int_{-a}^{x} r^{-1}dt = O(1) \quad \text{as} \quad x \to -a, \tag{7.1}$$

$$r(x)[y_1'(x)/y_1(x)]\int_{x}^{a} r^{-1}dt = O(1) \quad \text{as} \quad x \to a. \tag{7.2}$$

Let u *be absolutely continuous on* $[-a, a]$ *with* $u(-a)=u(a)$, $\int_{-a}^{a} ru'^2dx<\infty$, *and*

$$u(a)\int_{-a}^{a} su\, dx \leqq 0. \tag{7.3}$$

Then

$$\int_{-a}^{a} ru'^2dx \geqq \int_{-a}^{a} su^2dx + u^2(a)\int_{-a}^{a} sdx. \tag{7.4}$$

If $\int_{-a}^{a} ry_1'^2dx = \infty$, *or* $\alpha = \lim_{x\to -a} r(x)y_1'(x) = \infty$, *or* $\beta = \lim_{x\to a} r(x)y_1'(x) = -\infty$, *equality holds in* (7.4) *only for* $u \equiv 0$. *Otherwise, equality holds if and only if either* $u \equiv ky_1$, *or*

$$u(x) = u(a)\left\{1 - (\alpha - \beta)^{-1}\left(\int_{-a}^{x} sdx\right)y_1(x)\right\}. \tag{7.5}$$

Proof. We see that ry_1' is strictly decreasing on $(-a, a)$. Since $y_1'(\bar{x}) = 0$ for at least one $\bar{x} \in (-a, a)$, it follows, in fact, that ry_1' has precisely one zero on $(-a, a)$. Hence the limits α and β both exist (finite or infinite) with $\alpha > 0$, $\beta < 0$.

Now, if we replace (a, b) by $(-a, a)$ in (6.4)–(6.6) and take $y = y_1$ and $C = u(a) = u(-a)$, the conditions (7.1), (7.2) imply (using Schwarz's inequality) that $A = B = 0$, so that

$$\int_{-a}^{a} ru'^2dx \geqq \int_{-a}^{a} su^2dx + u^2(a)\int_{-a}^{a} sdx - 2u(a)\int_{-a}^{a} su\,dx.$$

Thus (7.4) follows from (7.3), and equality can hold in (7.4) only if $u(x) \equiv u(a) + ky_1(x)$ and equality holds in (7.3).

If $\int_a^b ry_1'^2dx = \infty$, we see (as in previous cases) that equality holds in (7.4) only if $k = u(a) = 0$, that is, only if $u \equiv 0$. Suppose that $u = u(a) + ky_1$. Then the integral

$$\int_{-a}^{a} s[u(a) + ky_1]dx = u(a)\int_{-a}^{a} sdx - k(\beta - \alpha)$$

exists if and only if α, β are both finite. Thus if α or β are infinite, equality can hold in (7.4) only if $u(x) \equiv u(a)$, which by (7.3) reduces to $u(x) \equiv 0$. If $\int_{-a}^{a} ru'^2dx < \infty$, and α, β are both finite, then $A = B = 0$ as noted above when $u = u(a) + ky_1$, and equality holds in (7.4) if and only if

$$u(a)\int_{-a}^{a} s[u(a) + ky_1]dx = u(a)\left\{u(a)\int_{-a}^{a} sdx - k(\beta - \alpha)\right\} = 0,$$

which reduces to the conditions quoted in the theorem.

REMARK 1. As usual, if $u(a) = u(-a) = 0$, the existence of the integrals $\int_{-a}^{a} sdx$ and $\int_{-a}^{a} su\,dx$ is not required. In this theorem, by taking

$$u(x) = \begin{cases} 0, & -a \leqq x \leqq a', \\ y_1(x) - y_1(a'), & a' \leqq x \leqq b', \qquad \textit{with } y_1(b') = y_1(a'), \\ 0, & b' \leqq x \leqq b, \end{cases}$$

one can show that *the inequality* (7.4) *is sharp if either*

$$\int_{-a}^{0} ry_1'^2dx = \infty, \qquad \varliminf_{x\to -a} ry_1y_1' < \infty, \quad \textit{and} \quad ry_1' = O(1) \quad \textit{as} \quad x \to a, \tag{7.6}$$

or

$$(7.7)\quad \int_0^a r y_1'^2 dx = \infty, \qquad \overline{\lim_{x\to a}}\, r y_1 y_1' > -\infty, \quad \textit{and} \quad r y_1' = O(1) \quad \textit{as} \quad x \to -a.$$

By taking $a=\pi/2$, $r\equiv s\equiv 1$, and $y_1=\cos x$ in this theorem, we see that *if* $u'\in L_2$, $u(-\pi/2)=u(\pi/2)$, *and* $u(\pi/2)\int_{-\pi/2}^{\pi/2} u\,dx \leqq 0$, *then*

$$(7.8)\qquad \int_{-\pi/2}^{\pi/2} u'^2 dx > \int_{-\pi/2}^{\pi/2} u^2 dx + \pi u^2(\pi/2),$$

unless $u=k\cos x$ *or* $u(x)=u(\pi/2)(1-(\pi/2)\cos x)$. By translating the interval of integration from $(-\pi/2, \pi/2)$ to $(0, \pi)$, we obtain Theorem 257 of [**29**]. By combining the translated version of (7.8) with (6.17), we obtain the inequality

$$(7.9)\qquad \int_0^\pi u''^2 dx > \int_0^\pi u^2 dx + \pi u^2(\pi/2) \quad \textit{unless} \quad u = k\cos x,$$

provided $u''\in L_2$, $u'(0)=u'(\pi)=0$, *and* $u(\pi/2)\int_0^\pi \cos x\; dx \leqq 0$. This includes a result proved by Fan, Taussky, and Todd [**27**]; see also [**4**, p. 477]. Theorem 7.1 is an extension of Theorem 1.3* of [**4**], and additional examples are given there, including (3.11) of the present paper.

From now on we shall deal only with the (symmetric) case that r and s are even functions on $(-a, a)$. For our next theorem we require the following lemma concerning solutions of equation (6.2) (cf. [**4**, Lemma 1.2]):

LEMMA 7.1. *Suppose r and s are both positive, continuous, and symmetric on $(-a, a)$, where $0<a\leqq\infty$. Let equation* (6.2) *have an even solution $y_1(x)$ and an odd solution $y_2(x)$ such that $y_1(\pm a)=0$, $y_1(x)>0$ for $-a<x<a$, $y_2'(x)>0$ for $0\leqq x<a$, and*

$$\gamma = \lim_{x\to a} r(x)[y_1'(x)/y_2(x)] \quad \textit{exists (finite)}.$$

Let y be any solution of (6.2) *which is not a multiple of y_1. Then*

$$(7.10)\qquad \beta_y \equiv \lim_{x\to a} r(x)[y'(x)/y(x)] \geqq \lim_{x\to -a} r(x)[y'(x)/y(x)] \equiv \alpha_y,$$

and α_y, β_y are both finite. Moreover, equality holds if and only if $\beta_{y_2}=0$.

Proof. We may write $y=Ay_1+By_2$, $y'=Ay_1'+By_2'$, where $B\neq 0$. Since $y_1(\pm a)=\lim_{x\to\pm a} y_1(x)=0$, we have

$$\beta_y = \lim_{x\to a}\frac{r(x)\{Ay_1'(x)+By_2'(x)\}}{Ay_1(x)+By_2(x)} = \beta_{y_2} + \frac{A}{B}\gamma,$$

$$\alpha_y = \lim_{x\to a}\frac{r(-x)y'(-x)}{y(-x)} = \lim_{x\to a}\frac{r(x)\{-Ay_1'(x)+By_2'(x)\}}{Ay_1(x)-By_2(x)} = -\beta_{y_2} + \frac{A}{B}\gamma,$$

provided β_{y_2} exists and is finite. Hence, the result will be proved if we can show that β_{y_2} exists and is nonnegative. However, this follows from the fact that

$$(ry_2'/y_2)' = \{y_2(ry_2')' - ry_2'^2\}/y_2^2 = -(sy_2^2 + ry_2'^2)/y_2^2 < 0$$

on the interval $[0, a)$, so that ry_2'/y_2 is a positive decreasing function on $(0, a)$.

REMARK 2. It was shown in [4] for the case $r(x) \equiv 1$ and s continuous on the *compact* interval $[-a, a]$, that if s is nonincreasing on $[0, a]$, then the existence of an even solution y_1 satisfying $y_1(\pm a) = 0$ and $y_1(x) > 0$ on $(-a, a)$ implies the existence of an odd solution y_2 with $y_2'(x) > 0$ on $[0, a)$. It was also shown in this case that $\beta_{y_2} = 0$ if and only if s is constant. In our case one can prove the same results in case the product rs is nonincreasing on the (not necessarily bounded) interval $[0, a)$. In the compact case considered in [4] there was no need to assume the existence of the limit γ, this existence being (almost) axiomatic. In our case, however, although ry_1' is decreasing and y_2 is increasing, the quotient ry_1'/y_2 is a negative decreasing function on $(0, a)$, so that γ could conceivably be $-\infty$.

THEOREM 7.2. *Suppose that r and s are positive, continuous, and even on $(-a, a)$, where $0 < a \leqq \infty$, and that equation* (6.2) *has even and odd solutions, y_1 and y_2, satisfying the hypotheses of Lemma* 7.1, *and*

$$r(x)[y_1'(x)/y_1(x)] \int_x^a r^{-1} dt = O(1) \qquad \text{as } x \to a. \tag{7.11}$$

Let u be absolutely continuous on $[-a, a]$ with $u(-a) = u(a)$ and $\int_{-a}^a ru'^2 dx < \infty$, and suppose that u vanishes for even a single point of $[-a, a]$. Then

$$\int_{-a}^a ru'^2 dx \geqq \int_{-a}^a su^2 dx. \tag{7.12}$$

If $\int_{-a}^a ry_1'^2 dx = \infty$, or if either of the limits $\alpha = \lim_{x\to -a} ry_1'$, $\beta = \lim_{x \to a} ry_1'$, is infinite, then equality holds in (7.12) *only for $u \equiv 0$; otherwise equality holds if and only if $u \equiv ky_1$.*

Proof. Suppose first that $u(x)$ vanishes for $x = a$, so $u(-a) = u(a) = 0$. By the symmetry of r and y_1, it follows from (7.11) that both the hypotheses (7.1), (7.2) of Theorem 7.1 are satisfied. Since $u(a) = 0$, the inequality (7.12) follows from Theorem 7.1 in this case, with the conditions for equality as noted.

Next, suppose that $u(\pm a) \neq 0$, so that $u(\bar{x}) = 0$, where $-a < \bar{x} < a$. In this case, let y be the unique solution of (6.2) determined by the initial conditions $y(\bar{x}) = 0$, $y'(\bar{x}) = 1$. By Sturm's separation theorem (cf. [**10**, Th. 1]), y has no other zeros on $(-a, a)$. Using this solution y, and $C = u(\bar{x}) = 0$ in (6.3)–(6.5), we have, using the notation of Lemma 7.1,

$$A = \varliminf_{x \to -a} r(x) \frac{y'(x)}{y(x)} u^2(x) = \alpha_y u^2(-a) = \alpha_y u^2(a),$$

$$B = \overline{\lim_{x\to a}}\, r(x) \frac{y'(x)}{y(x)} u^2(x) = \beta_y u^2(a).$$

Thus (6.4) reduces to

$$\int_{-a}^{a} ru'^2dx \geqq (\beta_y - \alpha_y)u^2(a) + \int_{-a}^{a} su^2dx \geqq \int_{-a}^{a} su^2dx,$$

by Lemma 7.1, proving (7.12).

In this case equality can hold in (7.12) only if $u=ky$ and $\beta_y=\alpha_y$. If now $k\neq 0$, then the condition $u(-a)=u(a)$ implies $y(-a)=y(a)$. Since $y=Ay_1+By_2$ with $B\neq 0$ and $y_1(\pm a)=0$, this condition reduces to $-By_2(a)=By_2(a)$. This is a contradiction since (see Lemma 7.1) y_2 is positive and increasing on $(0, a)$. Hence $k=0$, so equality can hold in (7.12) only for $u(x)\equiv 0$, which is impossible in the case under consideration. We conclude that equality is never attained in (7.12) for any admissible function u which is not identically zero but which has a zero at an interior point of $[-a, a]$. This proves slightly more than was asserted in the statement of the theorem.

By taking $a=\pi/2$, $r(x)\equiv s(x)\equiv 1$, $y_1=\cos x$, $y_2=\sin x$ (giving $\gamma=-1$, and $\beta_{y_2}=0$), we obtain the following result: *Let $u(-\pi/2)=u(\pi/2)$, where $u'\in L_2$ and u vanishes at least once on $[-\pi/2, \pi/2]$. Then*

$$\int_{-\pi/2}^{\pi/2} u'^2dx > \int_{-\pi/2}^{\pi/2} u^2dx \qquad \textit{unless } u = k\cos x. \tag{7.13}$$

A comparison of (7.13) with (7.8) shows the peculiar character of the last theorem.

8. Wirtinger type inequalities (III). The inequality usually called Wirtinger's inequality (see (8.10)) does not follow from any of the results so far. In order to obtain (an extension of) this inequality, we require an extension of the classical Sturm separation theorem for (6.2) to noncompact intervals. We require the following definition, in which we suppose that r and s are continuous, and r is positive, on an interval $[0, a)$ where $0<a\leqq\infty$. We say that a is *singular* for the equation (6.2) if either $a=-\infty$, or if a is finite but either of the limits $\lim_{x\to a} r(x)$ or $\lim_{x\to a} s(x)$ do not exist, or if $\lim_{x\to a} r(x)=0$. Otherwise a is said to be *nonsingular*; all other points of $[0, a)$ are also called *nonsingular*. If a is nonsingular for (6.2), then solutions y of (6.2) can clearly be extended to a so that y and y' are continuous at a. The following lemma is part of Theorem 1 proved in [**10**]:

LEMMA 8.1. *Let y_1, y_2 be two linearly independent solutions of* (6.2) *on* $[0, a)$. *Suppose that x_1 and $\bar{x}_1$, where $0\leqq x_1<\bar{x}_1\leqq a$, are consecutive zeros of y_1. Then* (i) *if $\bar{x}_1$ is nonsingular for* (6.2), *y_2 has precisely one zero on $(x_1, \bar{x}_1)$*, (ii) *if $\bar{x}_1$ is singular for* (6.2) *and if $y_2(x)\neq 0$ for $x_1<x<a$, we have $y_2(x)=O[y_1(x)]$ as $x\to a$, so that* $y_2(a)=0$.

Part (i) of this lemma is just the Sturm separation theorem. We shall also

need the following lemma, proved in [**10**, Th. 2] using a slightly different notation:

LEMMA 8.2. *Let r and s be continuous with r positive on $(-a, a)$, where $0<a\leqq\infty$, and suppose that equation* (6.2) *has a solution y_1 with consecutive zeros at $-a, 0, a$. Let u be a function which is continuous on $[-a, a]$, with $u(-a)=u(a)$. Then either every solution of* (6.2) *which vanishes at a point of $(-a, a)$ has only a single zero on $(-a, a)$, and zeros at both $-a$ and a, or there exists a solution y of* (6.2), *and two points x_1, x_2 with $-a\leqq x_1<0\leqq x_2<a$ (or $-a<x_1\leqq 0<x_2\leqq a$), such that $u(x_1)=u(x_2)$ and $y(x_1)=y(x_2)=0$, while $y(x)\neq 0$ for any other points of $(-a, a)$.*

LEMMA 8.3. *Suppose that r and s are both positive, continuous, and even on $(-a, a)$, where $0<a\leqq\infty$. Let equation* (6.2) *have an odd solution y_1 with consecutive zeros at $-a$, 0, a, and an even solution y_2 such that $r(x)y_2'(x)/y_2(x)$ is defined and nonnegative for all x sufficiently near a. Suppose also that the limit*

$$\gamma = \lim_{x\to a} r(x)[y_1'(x)/y_2(x)]$$

exists (finite), and that y is any solution of (6.2) *which is not a multiple of y_1. Then*

$$\beta_y \equiv \lim_{x\to a} r(x)[y'(x)/y(x)] \geqq \lim_{x\to -a} r(x)[y'(x)/y(x)] \equiv \alpha_y, \tag{8.1}$$

and α_y, β_y are both finite. Moreover, $y_2(a)=\lim_{x\to a} y_2(x)$ exists, with $0<|y_2(a)|\leqq\infty$, and y_2 has precisely one zero on $(0, a)$. Finally, equality holds in (8.1) *if and only if $\beta_{y_2}=0$, and this is the case if $\beta=\lim_{x\to a} r(x)y_2'(x)=0$.*

Proof. The proof of the main part of the lemma is essentially the same as the proof of Lemma 7.1 once we note that $\lim_{x\to a} y_1(x)/y_2(x)=0$ necessarily holds. This follows from the fact that $\lim y_1(x)=y_1(a)=0$, while for all x sufficiently close to a, $y_2'(x)$ and $y_2(x)$ have the same sign. That is, if $y_2(x)>0$, then $y_2'(x)\geqq 0$, so y_2 is positive and increasing with $\lim y_2(x)>0$; similarly, $\lim y_2(x)<0$ if $y_2(x)<0$ for x near a. This also proves that $0<|y_2(a)|\leqq\infty$. The fact that y_2 has precisely one zero on $(0, a)$ follows from the Sturm separation theorem if a is nonsingular (Lemma 8.1(i)). If a is singular and y_2 had no zeros on $(0, a)$, then by Lemma 8.1(ii), we'd have $y_2(a)=0$, which is not the case. If $\beta=0$, then $\beta_{y_2}=0$ follows from the fact that $y_2(a)\neq 0$.

The last two lemmas were proved in the compact case in [**4**, Lemmas 1.4, 1.5] where the proof was somewhat easier.

THEOREM 8.1. *Suppose that r and s are positive, continuous, and even on $(-a, a)$, where $0<a\leqq\infty$, and that equation* (6.2) *has odd and even solutions, y_1 and y_2, satisfying the hypotheses of Lemma* 8.3. *In addition, we assume that $\int_{-a}^{a} s\,dx<\infty$, and that*

$$r(x)[y_1'(x)/y_1(x)]\int_x^a r^{-1}dt = O(1) \qquad \textit{as } x\to a. \tag{8.2}$$

Let u be absolutely continuous on $[-a, a]$ with $u(-a)=u(a)$ and $\int_{-a}^{a} ru'^2dx<\infty$, and suppose that u satisfies the orthogonality condition

$$\int_{-a}^{a} su\, dx = 0. \tag{8.3}$$

Then

$$\int_{-a}^{a} ru'^2dx \geqq \int_{-a}^{a} su^2dx. \tag{8.4}$$

Moreover, equality holds in (8.4) *if and only if $u=A_1y_1+A_2y_2$, where $A_1=0$ if $\int_{-a}^{a} ry_1'^2dx=\infty$ or $\alpha=\lim_{x\to a} r(x)y_1'(x)$ is infinite, and $A_2=0$ if $\int_{-a}^{a} ry_2'^2dx=\infty$ or $\beta=\lim_{x\to a} r(x)y_2'(x)\neq 0$. (In case $\int_{-a}^{a} ry_2'^2dx<\infty$, α is finite; see also the following remark.)*

Proof. We begin the proof by noting that if $\int_{-a}^{a} ry_2'^2dx<\infty$, then (8.2) implies that $y_2(a)$ is finite, that α is finite, and that $\beta_{y_2}=0$ if and only if $\beta=0$. That $y_2(a)=\lim_{x\to a} y_2(x)$ exists and is finite follows from Cauchy's limit criterion, and

$$|\, y_2(x) - y_2(x')\,| = \left| \int_{x'}^{x} r^{-1/2}r^{1/2}y_2'dt \right| \leqq \left(\int_{x'}^{x} r^{-1}dt \right)^{1/2} \left(\int_{x'}^{x} ry_2'^2\, dt \right)^{1/2},$$

together with the fact that (8.2) implies $\int_x^a r^{-1}dt<\infty$. Since $\gamma=\lim_{x\to a}(ry_1')/y_2$ exists (finite), α must also be finite, and $\beta_{y_2}=\lim_{x\to a}(ry_2')/y_2=0$ if and only if $\beta=0$, since $0<|y_2(a)|<\infty$.

Now since y_2 is even, it has precisely two zeros on $(-a, a)$ by Lemma 8.3. It follows that the first alternative in the conclusion of Lemma 8.2 does not occur here. Using the remaining alternative, we conclude that there exists a solution y of (6.2) and two points x_1, x_2, with $-a\leqq x_1<0\leqq x_2<a$ (or $-a<x_1\leqq 0<x_2\leqq a$), such that $u(x_1)=u(x_2)$ and $y(x_1)=y(x_2)=0$, while $y(x)\neq 0$ for any other points of $(-a, a)$. Using this solution y in (6.3)–(6.5), and setting $C=u(x_1)=u(x_2)$—and replacing (a, b) by $(-a, a)$—we see that

$$\begin{cases} A = \varliminf_{x\to -a} r(x)[y'(x)/y(x)]\{u(x)-u(x_1)\}^2 = \alpha_y\{u(a)-u(x_1)\}^2, \\ B = \varlimsup_{x\to a} r(x)[y'(x)/y(x)]\{u(x)-u(x_1)\}^2 = \beta_y\{u(a)-u(x_1)\}^2, \end{cases} \tag{8.5}$$

are both finite by Lemma 8.3, if $-a\leqq x_1<0<x_2<a$ (or $-a<x_1<0<x_2\leqq a$) so that y is not a multiple of y_1. Moreover, in this case (as can be seen by the analysis following (6.5)), the inequality (6.6) remains valid and reduces to

$$\int_{-a}^{a} ru'^2dx \geqq \int_{-a}^{a} su^2dx + u^2(x_1)\int_{-a}^{a} sdx + B - A. \tag{8.6}$$

If $x_1=0$ (so $x_2=a$) or $x_2=0$ (so $x_1=-a$), then y is necessarily a multiple of y_1, and y has precisely three zeros on $[-a, a]$ at $-a$, 0, a, and $u(-a)=u(0)=u(a)$. Indeed, *whenever $u(a)=u(0)$ we may (and do) always take $y=y_1$*. In this

case, $C=u(x_1)=u(x_2)=u(0)$ and, by (8.2), it follows as in previous cases that $A=B=0$ and that (8.4) is valid, with equality only if $u(x)=u(0)+ky_1(x)$ and $u(x_1)=u(0)=0$, that is only if $u\equiv ky_1$. It remains to determine whether $u=ky_1$ is admissible for $k\neq 0$. The requirement $u(-a)=u(a)$ clearly holds. Also,

$$\int_{x'}^{x} sy_1 dt = -\int_{x'}^{x} (ry_1')' dt = r(x')y_1'(x') - r(x)y_1'(x). \tag{8.7}$$

Hence, $\int_{-a}^{a} sy_1 dt$ exists if and only if $\alpha=\lim_{x\to a} r(x)y_1'(x)$ exists (finite). (In this case of course $\int_{-a}^{a} sy_1 dt=0$, since the integrand is an odd function.) We see that $u=ky_1$ is admissible for $k\neq 0$ if and only if $\int_{-a}^{a} ry_1'^2 dx<\infty$ and α is finite. Hence, in the case we are considering ($u(0)=u(a)$), equality holds in (8.4) if and only if $u\equiv ky_1$, where $k=0$ unless both $\int_{-a}^{a} ry_1'^2 dx<\infty$ and α is finite.

If $u(0)\neq u(a)$, then neither x_1 nor x_2 is 0, so y is not a multiple of y_1. It now follows from (8.5), (8.6), and Lemma 8.3 that (8.4) is valid, and that equality can hold only if $u(x)=u(x_1)+ky(x)$, $u(x_1)=0$, and either $\beta_{y_2}=0$ or $u(a)=u(x_1)$, thus only if $u\equiv ky$ and either $\beta_{y_2}=0$ or $u(-a)=u(a)=u(x_1)=0$. We may write $y=Ay_1+By_2$ where $B\neq 0$, so that the necessary conditions for equality in (8.4) become

$$u = kAy_1 + kBy_2 \quad (B \neq 0), \quad \text{and either } \beta_{y_2} = 0 \quad \text{or } u(-a) = u(a) = 0. \tag{8.8}$$

In order to determine the admissibility of such functions u, we first observe that if $\int_{-a}^{a} ry_1'^2 dx<\infty$ and $\int_{-a}^{a} ry_2'^2 dx<\infty$, then $\int_{-a}^{a} ry_1' y_2' dx$ exists (by the Cauchy-Schwarz inequality), and this integral has the value zero since the integrand is odd. In this case, with u given by (8.8), we have

$$\int_{-a}^{a} ru'^2 dx = k^2A^2\int_{-a}^{a} ry_1'^2 dx + k^2B^2\int_{-a}^{a} ry_2'^2 dx, \tag{8.9}$$

so that u is admissible if we also have $\int_{-a}^{a} sudt=0$ and $u(-a)=u(a)$. If, however, $\int_{-a}^{a} ry_2'^2 dx=\infty$ and $k\neq 0$, one sees that $\int_{-a}^{a} ru'^2 dx=\infty$; similarly, if $\int_{-a}^{a} ry_1'^2 dx=\infty$ and $kA\neq 0$, then u is not admissible. As noted above, $u=ky_1$ is admissible if $\int_{-a}^{a} ry_1'^2 dx<\infty$ and α is finite. For y_2 we proceed as at (8.7) and see that $\int_{-a}^{a} sy_2 dt$ exists if and only if $\beta=\lim_{x\to a} r(x)y_2'(x)$ is finite, and that $\int_{-a}^{a} sy_2 dt=0$ if and only if $\beta=0$. Thus ky_2 is admissible if and only if $\int_{-a}^{a} ry_2'^2 dx<\infty$ and $\beta=0$; note that the first of these conditions implies that $y_2(a)$ is finite, hence that $y_2(-a)=y_2(a)$ since y_2 is even. Summarizing, we have now shown that:

(a) $\quad u = ky_1$ is admissible for $k \neq 0 \Leftrightarrow \displaystyle\int_{-a}^{a} ry_1'^2 dx < \infty$ and α is finite;

(b) $\quad u = ky_2$ is admissible for $k \neq 0 \Leftrightarrow \displaystyle\int_{-a}^{a} ry_2'^2 dx < \infty$ and $\beta = 0$;

(c) $\quad$ if $\displaystyle\int_{-a}^{a} ry_2'^2 dx < \infty$. then α is finite, and $\beta_{y_2} = 0 \Leftrightarrow \beta = 0$.

To complete the discussion of the cases of equality in (8.4), *assume first that either* $\int_{-a}^{a} ry_2'^2dx = \infty$ *or* $\beta \neq 0$. Then by (b), ky_2 is not admissible unless $k=0$, so by (8.8) it follows that equality holds in (8.4) if and only if $u\equiv 0$. Next, *suppose that* $\int_{-a}^{a} ry_2'^2dx < \infty$ *and* $\beta = 0$. By (b), ky_2 is admissible, and from (a) and (c) it follows that ky_1 is admissible (for $k\neq 0$) if and only if $\int_{-a}^{a} ry_1'^2dx < \infty$. Hence, since $\beta_{y_2}=0$ by (c), it follows from (8.8) that equality holds in (8.4) if and only if $u = A_1y_1 + A_2y_2$, where $A_1 = 0$ if $\int_{-a}^{a} ry_1'^2dx = \infty$. By using (c), a (tedious) logical analysis of the conditions for equality in (8.4) obtained in the two cases ($u(0) = u(a)$ and $u(0) \neq u(a)$) shows that they may be combined into the form given in the statement of the theorem.

REMARK. By using Abel's identity [**10**, Th. 1] for y_1 and y_2, one can easily show that the limit α is *always* finite under our hypotheses. This implies that $\alpha_1 = \lim_{x\to a} ry_1y_1' = 0$ always, and that equality holds in (8.4) for $u\equiv ky_1$ provided $\int_{-a}^{a} ry_1'^2dx < \infty$. Using these facts, one easily proves that (8.4) is always sharp, even if $\int_{-a}^{a} ry_1'^2dx = \infty$. To see this, let g be any positive function on $[0, a)$ such that $g(a) = 0$ and $\int_0^a rg'^2dx < \infty$, chosen so that $g(a') = y_1(a')$ for a' "sufficiently close" to a, and define u on $[0, a]$ by

$$u(x) = \begin{cases} y_1(x), & 0 \leqq x \leqq a', \\ g(x), & a' \leqq x \leqq a. \end{cases}$$

Set $u(-x) = -u(x)$ for $x\in[0, a]$, so u is an odd, admissible function. We then find that, given $\delta\in(0, 1)$, we have

$$\int_0^a su^2dx > (1-\delta)\int_0^a ru'^2dx,$$

provided $r(a')y_1(a')y_1'(a') + (1-\delta)\int_{a'}^{a} rg'^2dx < \delta\int_0^{a'} ry_1'^2dx$.

Wirtinger's inequality is the special case of this theorem obtained by taking $a=\pi$, $r\equiv s\equiv 1$, $y_1 = \sin x$, $y_2 = \cos x$. One finds that $\alpha = -1$, $\beta = 0$, and $\gamma = 1$. Hence,

(8.10) *if* $u' \in L_2$, $u(-\pi) = u(\pi)$, *and* $\int_{-\pi}^{\pi} udx = 0$, *then*

$$\int_{-\pi}^{\pi} u'^2dx > \int_{-\pi}^{\pi} u^2dx \textit{ unless } u = A_1\sin x + A_2\cos x.$$

An example for the noncompact case is given by taking $a=\infty$, $r\equiv 1+x^2$, $s\equiv 4(1+x^2)^{-1}$, $y_1 = x(1+x^2)^{-1}$, $y_2 = (1-x^2)(1+x^2)^{-1}$, for which $\alpha = -1$, $\beta = 0$, and $\gamma = 1$ again. This gives

(8.11) *if* $u(-\infty) = u(\infty)$, *and* $\int_{-\infty}^{\infty} u(1+x^2)^{-1}dx = 0$, *then*

$$\int_{-\infty}^{\infty}(1+x^2)u'^2dx \geqq 4\int_{-\infty}^{\infty}\frac{u^2}{1+x^2}\,dx,$$

with equality if and only if $u=\{Ax+B(1-x^2)\}(1+x^2)^{-1}$.

The best-known application of Wirtinger's inequality is to the proof of the isoperimetric inequality for plane curves, although other geometric inequalities are given in Blaschke [**12**, pp. 105–109]. (Incidentally, as pointed out in [**42**], the name Wirtinger appears to have first been attached to the inequality (8.10) in [**12**, p. 105], under the heading "Ein Lemma von Wirtinger." It is not clear whether—or when—Wirtinger ever published his own proof. Blaschke gives no reference but merely the acknowledgement "Den folgenden noch durchsichtigeren Beweis verdanke ich Herrn W. Wirtinger.") We shall give a proof of the isoperimetric inequality following Janet [**34**], differing only slightly from the short proof given in [**29**, pp. 186–187]. To this end, we first transform the inequality (8.10) from $(-\pi,\pi)$ to $(0,L)$, by setting $x=-\pi+(2\pi t/L)$, to obtain the result that *if* $U'\in L_2$ *and* $U(0)=U(L)$, $\int_0^L U\,dt=0$, *then*

$$\int_0^L U'^2dt > \frac{4\pi^2}{L^2}\int_0^L U^2dt \text{ unless } U = A_1\sin\frac{2\pi t}{L} + A_2\cos\frac{2\pi t}{L}. \tag{8.12}$$

Now, let y be any function such that $y'\in L_2$ and $y(0)=y(L)$. Setting $U(s)=y(s)-(\int_0^L y\,dt/L)$, we have $U'\in L_2$, $U(0)=U(L)$, and $\int_0^L U\,ds=0$, so that on writing $B=(\int_0^L y\,ds/L)$, we have

$$\int_0^L y'^2ds > \frac{4\pi^2}{L^2}\int_0^L (y-B)^2ds \text{ unless } y-B = A_1\sin(2\pi s/L) + A_2\cos(2\pi s/L). \tag{8.13}$$

With these analytic preliminaries disposed of, let C be any simple, closed, positively oriented curve given by the parametric equations $x=x(s)$, $y=y(s)$, $0\leqq s\leqq L$, where s denotes arc length (so L is the length of C), and we assume that $x', y'\in L_2$. We have $x(0)=x(L)$ and $y(0)=y(L)$, and if A denotes the area of the region bounded by C, then $A=-\int_C y\,dx=-\int_0^L yx'ds$. Adding corresponding sides of the inequalities (8.13) and using

$$\int_0^L\left\{x'+\frac{2\pi}{L}(y-B)\right\}^2ds > 0 \text{ unless } x' = -\frac{2\pi}{L}(y-B),$$

we obtain

$$\int_0^L (x'^2+y'^2)ds + \frac{4\pi}{L}\int_0^L (y-B)x'ds > 0.$$

Using $x(0)=x(L)$ and $x'^2+y'^2\equiv 1$, we thus have $L^2>-4\pi\int_0^L yx'ds=4\pi A$, unless both $y-B=A_1\sin(2\pi s/L)+A_2\cos(2\pi s/L)$, and $x-K=A_1\cos(2\pi s/L)-A_2\sin(2\pi s/L)$. That is, $L^2>4\pi A$ holds unless the curve C is a circle: $(x-K)^2+(y-B)^2=A_1^2+A_2^2$.

9. Other inequalities of Wirtinger type. In this final section we shall merely

list some of the more interesting inequalities which bear a family resemblance to Wirtinger's inequality, and mention other related results. No proofs will be given, and no attempt at completeness is made here.

Of those inequalities involving only first derivatives, one of the most interesting is the following result of E. Schmidt [51]. *Let u be absolutely continuous on $[0, L]$ with $u(0)=u(L)$, and let m and M be the minimum and maximum values of u on $[0, L]$. Then*

$$\int_0^L \left| u(t) - \tfrac{1}{2}(m+M) \right|^b dt \leqq \frac{1}{b-1}\left(\frac{bL}{4\pi}\sin\frac{\pi}{b}\right)^b \int_0^L |u'(t)|^b dt, \tag{9.1}$$

for $b \geqq 1$. By taking $b=2$, and adding the condition $\int_0^L u\,dt=0$, we obtain

$$\int_0^L u'^2 dt \geqq \frac{4\pi^2}{L^2}\int_0^L u^2 dt + \frac{\pi^2}{L}(m+M)^2, \tag{9.2}$$

an improvement of Wirtinger's inequality (8.12). The inequality (9.2) was obtained again recently by Benson [11], and used by him to give an interesting improvement of the isoperimetric inequality. Schmidt's result (9.1) also includes as a special case (with $b=2k$, an even integer) the inequality

$$\int_{-\pi}^{\pi} u^{2k} dx \leqq \frac{1}{2k-1}\left(k \sin\frac{\pi}{2k}\right)^{2k} \int_{-\pi}^{\pi} u'^{2k} dx, \tag{9.3}$$

proved by Beesack [5] for functions u with $u' \in L_{2k}$, $u(-\pi)=u(\pi)$, and satisfying $\int_{-\pi}^{\pi} u^{2k-1}dx=0$. Inequalities similar to those in Section 8 involving positive functions r, s and a side condition of the form $\int su^{p-1}dx=0$ are also given in [5].

Diaz and Metcalf [26] proved a large number of very general inequalities of Wirtinger type (by quite elementary methods), of which the following is typical: *Let $u \in C^1[a, b]$, and let t_1, t_2 be real numbers such that $a \leqq t_1 \leqq t_2 \leqq b$. Suppose that $u(a)=u(b)$, $u(t_1)=u(t_2)$, and $(b-a)^2u^2(t_1) \geqq 2u(t_1)\int_a^b u\,dx$. Then*

$$\int_a^b u^2 dx \leqq \frac{1}{\pi^2}\max\left\{(t_2-t_1)^2, (b-a-t_2+t_1)^2\right\}\int_a^b u'^2 dx. \tag{9.4}$$

Wirtinger's inequality follows from (9.4) by taking $t_1=a$, $t_2=b$.

One of the earliest papers dealing with derivatives of higher order is that of M. Janet [34]. He proved that *if u is any function such that $u^{(k)}(a)=0=u^{(k)}(b)$ for $k=0, \cdots, n-1$, and $u^{(n)} \in C[a, b]$, then*

$$\int_a^b \left(u^{(n-1)}\right)^2 dx \leqq (b-a)^2 \omega_n^{-2} \int_a^b \left(u^{(n)}\right)^2 dx, \tag{9.5}$$

where ω_n is the smallest positive zero of the $n \times n$ determinant $\Delta = \| y_i^{(k)} \|$, the y_i being n linearly independent solutions of the equation $y^{(2n)}+y^{(2n-2)}=0$ satisfying specified initial conditions. The more general problem, with $u^{(n-1)}$ replaced by $u^{(p)}$ (for $0 \leqq p < n$) on the left side of (9.5), was also dealt with later by

Janet [35]. The special case $p=0$ of this problem was generalized by Bellman in [2] who proved that *if* $u^{(n)}\in L_{2k}$, $u(x+2\pi)\equiv u(x)$, *and* $\int_{-\pi}^{\pi} udx=0$, *then*

$$\int_{-\pi}^{\pi} u^{2k}dx \leqq a_n^{2k}\int_{-\pi}^{\pi}(u^{(n)})^{2k}dx, \tag{9.6}$$

where k, n *are positive integers and the* a_n *are certain constants* (not sharp). In the case $k=1$, $n=2$, sharp results of this general character were obtained by Fan, Taussky, and Todd [27], and Beesack [4, Th. 2.6]; general results similar to those proved in Section 8 were also obtained in [4] for the case $k=1$, $n=2$. (See also Benson [11, pp. 305–306].) By taking appropriate limits of discrete inequalities, Pfeffer [48, Cor. 1.1] obtained the best possible constants ($a_n\equiv 1$) in (9.6) for the special case $k=1$, under (essentially) the same hypotheses as Bellman. Moreover, by this technique, Pfeffer obtained sharp inequalities of the form

$$\int_a^b (u^{(p)})^2dx \leqq \alpha\int_a^b (u^{(n)})^2dx + A_{p,n}(\alpha)\int_a^b u^2dx, \quad (1\leqq p<n), \tag{9.7}$$

valid for all $u\in C^n[a,b]$, with $u^{(k)}(a)=u^{(k)}(b)$ for $0\leqq k\leqq n-1$, and $\alpha>0$. If the Wirtinger side condition $\int_a^b udx=0$ is added, the factor $A_{p,n}(\alpha)$ must be altered for $\alpha>[(b-a)/2\pi]^{2(n-k)}$.

We conclude by mentioning two inequalities of a more general kind. W. J. Coles [21] dealt with inequalities of the form

$$(-1)^n\int_a^b p(x)u^2(x)dx \leqq \int_a^b \{u^{(n)}(x)\}^2dx,$$

related to differential equations $y^{(2n)}-py=0$ having (appropriate) solutions y such that $(-1)^n p(x)y(x)\geqq 0$. Levin and Stečkin [38] dealt with the following very general problem which has points of contact with Brink [19]. Let A_n denote the class of all functions u such that $u^{(k)}(0)=u^{(k)}(2\pi)$ for $0\leqq k\leqq n-1$, $u^{(n-1)}$ is absolutely continuous, and $\int_0^{2\pi} udx=0$. For $0<p\leqq\infty$, let $\|u\|_p$ denote the usual L_p norm of u. Suppose that $1\leqq p\leqq q\leqq\infty$ and $\nu^{-1}=1+q^{-1}-p^{-1}$ (where $a^{-1}=0$ or ∞ if and only if $a=\infty$ or 0). If $u\in A_n$, then

$$\|u\|_q \leqq C_{n,\nu}\|u^{(n)}\|_p,$$

where $C_{n,\nu}=\min_\xi\|\varphi_n-\xi\|_\nu$ and $\varphi_n(t)=\pi^{-1}\sum_{k=1}^{\infty}k^{-n}\cos(kt-(n\pi/2))$. This theorem is sharp only when $q=\infty$.

References

1. E. F. Beckenbach and R. Bellman, Inequalities, 2nd rev. printing, Springer, Berlin, 1965.
2. R. Bellman, A note on periodic functions and their derivatives, J. London Math. Soc., 18 (1943) 140–142.
3. ———, A note on an inequality of E. Schmidt, Bull. Amer. Math. Soc., 50 (1944) 734–736.
4. P. R. Beesack, Integral inequalities of the Wirtinger type, Duke Math. J., 25 (1958) 477–498.

5. P. R. Beesack, Extensions of Wirtinger's inequality, Trans. Roy. Soc. Canada, 53 (3) (1959) 21–30.

6. ———, Elementary proofs of the extremal properties of the eigenvalues of the Sturm-Liouville equation, Canad. Math. Bull., 3 (1960) 59–77.

7. ———, Hardy's inequality and its extensions, Pacific J. Math., 11 (1961) 39–61.

8. ———, On an integral inequality of Z. Opial, Trans. Amer. Math. Soc., 104 (1962) 470–475.

9. ———, On certain discrete inequalities involving partial sums, Canad. J. Math., 21 (1969) 222–234.

10. ———, On Sturm's separation theorem (to appear in Canad. Math. Bull.).

11. D. C. Benson, Inequalities involving integrals of functions and their derivatives, J. Math. Anal. Appl., 17 (1967) 292–308.

12. W. Blaschke, Kreis und Kugel, Chelsea reprint, New York, 1949.

13. H. D. Block, A class of inequalities, Proc. Amer. Math. Soc., 8 (1957) 844–851.

14. ———, Discrete analogues of certain integral inequalities, *ibid.* 852–859.

15. D. W. Boyd and J. S. W. Wong, An extension of Opial's inequality, J. Math. Anal. Appl., 19 (1967) 100–102.

16. D. W. Boyd, Best constants in inequalities related to Opial's inequality, J. Math. Anal. Appl., 25 (1969) 378–387.

17. ———, Best constants in a class of integral inequalities, Pacific J. Math., 30 (1969) 367–383.

18. ———, On the exponent of an osculatory packing, Canad. J. Math., 23 (1971) 355–363.

19. J. Brink, Inequalities involving $\|f\|_p$ and $\|f^{(n)}\|_q$ for f with n zeros, Ph.D. Dissertation, Iowa State University, 1970.

20. J. Calvert, Some generalizations of Opial's inequality, Proc. Amer. Math. Soc., 18 (1967) 72–75.

21. W. J. Coles, A general Wirtinger type inequality, Duke Math. J., 27 (1960) 133–138.

22. ———, Wirtinger-type integral inequalities, Pacific J. Math., 11 (1960) 871–877.

23. R. Courant and D. Hilbert, Methods of Mathematical Physics, Vol. 1, Wiley, New York, 1953.

24. K. M. Das and P. R. Beesack, Extensions of Opial's inequality, Pacific J. Math., 26 (1968) 215–232.

25. K. M. Das, An inequality similar to Opial's inequality, Proc. Amer. Math. Soc., 22 (1969) 258–261.

26. J. B. Diaz and F. T. Metcalf, Variations of Wirtinger's inequality, in Inequalities (Proc. Symposium Wright-Patterson Air Force Base, Ohio, 1965), pp. 79–103, New York, 1967.

27. K. Fan, O. Taussky, and J. Todd, Discrete analogues of inequalities of Wirtinger, Monatsh. Math., 59 (1955) 73–90.

28. H. Flanders, Remarks on Almgren's interior regularity theorems, Illinois J. Math., 13 (1969) 707–716.

29. G. H. Hardy, J. E. Littlewood and G. Pólya, Inequalities, 2nd ed., Cambridge, 1952.

30. J. M. Holt, Integral inequalities related to non-oscillation theorems for differential equations, SIAM J., 13 (1965) 767–794.

31. L. K. Hua, On an inequality of Opial, Scientia Sinica, 14 (1965) 789–790.

32. R. A. Hunt, An extension of the Marcinkiewicz interpolation theorem to Lorentz spaces, Bull. Amer. Math. Soc., 70 (1964) 803–807.

33. R. A. Hunt and G. Weiss, The Marcinkiewicz interpolation theorem, Proc. Amer. Math. Soc., 15 (1964) 996–998.

34. M. Janet, Sur la méthode de Legendre-Jacobi-Clebsch et quelques-unes de ses applications, Bull. Sci. Math., (2) 53 (1929) 144–160.

35. ———, Détermination explicite de certains minima dans des problèmes sans conditions aux limites, C. R. Acad. Sci. Paris, 194 (1932), 2109–2111.

36. W. J. Kim, On the zeros of solutions of $y^{(n)}+py=0$, J. Math. Anal. Appl., 25 (1969) 189–208.

37. C. M. Lee, On a discrete analogue of inequalities of Opial and Yang, Canad. Math. Bull. 11 (1968) 73–77.

38. V. I. Levin and S. B. Stečkin, Inequalities, Amer. Math. Soc. Transl. (2) 14 (1960) 1–29. See, especially, pp. 11–13.

39. N. Levinson, On an inequality of Opial and Beesack, Proc. Amer. Math. Soc., 15 (1964) 565–566.

40. C. L. Mallows, An even simpler proof of Opial's inequality, Proc. Amer. Math. Soc., 16 (1965) 173.

41. P. M. Maroni, Sur l'inégalité d'Opial-Beesack, C. R. Acad. Sci. Paris Sér. A 264 (1967) 62–64.

42. D. S. Mitrinović and P. M. Vasič, An inequality ascribed to Wirtinger, and its variations and generalizations, Univ. Beograd. Publ. Elektrotehn. Fak. Ser. Mat. Fiz. No. 247–No. 273 (1969) 157–170.

43. ———, Analytic Inequalities, Springer, Berlin, 1970.

44. Z. Nehari, The Schwarzian derivative and schlicht functions, Bull. Amer. Math. Soc., 55 (1949) 545–551.

45. C. Olech, A simple proof of a certain result of Z. Opial, Ann. Polon. Math., 8 (1960) 61–63.

46. Z. Opial, Sur une inégalité, *ibid.*, 29–32.

47. R. N. Pederson, On an inequality of Opial, Beesack and Levinson, Proc. Amer. Math. Soc., 16 (1965) 174.

48. A. M. Pfeffer, On certain discrete inequalities and their continuous analogs, J. Res. Nat. Bur. Stand. B. Math. and Math. Phys., 70 B (1966) 221–231.

49. A. Pleijel, An inequality (Swedish), Mat. Tidsskr. A (1949) 67–69.

50. R. Redheffer, Inequalities with three functions, J. Math. Anal. Appl., 16 (1966) 219–242.

51. E. Schmidt, Über die Ungleichung, welche die Integrale über eine Potenz einer Funktion und über eine andere Potenz ihrer Ableitung verbindet, Math. Ann., 117 (1940) 301–326.

52. D. T. T. Shum, On integral inequalities related to Hardy's, to appear in Canad. Math. Bull.

53. G. Talenti, Una diseguaglianza integrale, Boll. Un. mat. Ital. (3) 21 (1966) 25–34.

54. ———, Sopra una diseguaglianza integrale, Ann. Scuola Norm. Sup. Pisa (3) 21 (1967) 167–188.

55. G. Tomaselli, A class of inequalities, Boll. Un. mat. Ital. IV Ser. 2 (1969) 622–631.

56. D. Willett, The existence-uniqueness theorem for an nth order linear ordinary differential equation, this MONTHLY, 75 (1968) 174–178.

57. J. S. W. Wong, A discrete analogue of Opial's inequality, Canad. Math. Bull., 10 (1967) 115–118.

58. G. S. Yang, On a certain result of Z. Opial, Proc. Japan Acad., 42 (1966) 78–83.

NIL SUBRINGS IN FINITENESS CONDITIONS

ROBERT C. SHOCK, Southern Illinois University

Throughout this paper R will always denote an associative ring. We say that an element x is **nilpotent** in R if $x^m=0$ for some integer m. A subring S is **nilpotent** in R if there is an integer n such that $a_1 \cdots a_n=0$ for all $a_1, \cdots, a_n$ in S. A subring N is **nil** in R if each element is nilpotent in N. A nil subring, however, need not be nilpotent. This paper investigates nil subrings in rings satisfying certain finiteness conditions. These finiteness conditions force nil subrings to be nilpotent. The paper is self-contained. It could serve as supplementary

material for a senior level course in ring theory or for a first year graduate course in algebra.

1. Finiteness conditions. We begin with some basic definitions. A ring R satisfies the **maximum condition** on a collection $\mathfrak{A}$ of right ideals if each increasing sequence of right ideals becomes constant in $\mathfrak{A}$. This means that for each sequence $A_1 \subseteq A_2 \subseteq \cdots$ in $\mathfrak{A}$ we have $A_n = A_{n+1} = \cdots$ for some n. A ring R satisfies the **minimum condition** on a collection $\mathfrak{A}$ of right ideals if each descending sequence of right ideals becomes constant in $\mathfrak{A}$. A ring R is **right Artinian** if R satisfies the minimum condition on the collection of all right ideals. A ring R is **right Noetherian** if R satisfies the maximum condition on the collection of all right ideals. A right Artinian ring with unity is a right Noetherian ring [**10**, p. 69].

For a nonempty subset S of R let $\mathfrak{r}(S) = \{x \in R : sx = 0 \text{ for all } s \text{ in } S\}$. This set is a right ideal in R. We call $\mathfrak{r}(S)$ the **right annihilator** of S and term a right ideal K of R a right annihilator if $K = \mathfrak{r}(S')$ for some appropriate subset S' of R. We similarly define the left annihilator $\mathfrak{l}(S)$ of S and term a left ideal as a left annihilator. We speak of R as satisfying the maximum condition on right annihilators if the collection of right annihilators satisfies the maximum condition. We similarly speak of R as satisfying the maximum condition on left annihilators, the minimum condition on right annihilators, and the minimum condition on left annihilators.

For x in R we write $\mathfrak{r}(x)$ instead of $\mathfrak{r}(\{x\})$. We write "xR" instead of "the principal right ideal generated by x." A right ideal K is called **essential** if K has a nonzero intersection with each nonzero right ideal of R. In 1951 R. E. Johnson considered the set of elements x in R such that $\mathfrak{r}(x)$ is essential [**9**]. We shall show later that this set is an ideal and is called the right **singular ideal** of R. Next we term those rings which contain no infinite direct sum of nonzero right ideals as right **finite dimensional rings**. It can be shown that for a right finite dimensional ring R there is an integer n such that R contains a direct sum of n-summands and the number of summands of any other direct sum of R is at most n. This unique number n is called the **dimension** of R, and we write $\dim R = n$. We need one more definition. If a ring is right finite dimensional and satisfies the maximum condition on right annihilators, then it is called a right **Goldie ring**. A right Noetherian ring is a right Goldie ring.

2. A summary of the results. When is a nil ring nilpotent? When is a nil ideal nilpotent? These are old but basic questions in ring theory. We cite some answers. *A nil ideal is nilpotent in a right Artinian ring* (Hopkins, 1939). *A nil ideal is nilpotent in a right Noetherian ring* (Levitzki, 1950). Several years later similar results were extended to cover nil subrings. *A nil subring is nilpotent in a right Artinian ring.* This is implied by Jacobson's generalization of Engel's Theorem ([**7**], p. 201, 1952). *A nil subring is nilpotent in a right Goldie ring* (Lanski, 1969). *If R satisfies the maximum condition on right annihilators and on left annihilators, then a nil subring is nilpotent* (Herstein and Small, 1964). Certain

annihilating properties characterize nilpotent rings. *A nil ring N is nilpotent if and only if the following three increasing sequences become constant*:

$$\mathfrak{r}(N) \subseteq \mathfrak{r}(N^2) \subseteq \mathfrak{r}(N^3) \subseteq \cdots,$$
$$\mathfrak{r}(x_1) \subseteq \mathfrak{r}(x_2x_1) \subseteq \mathfrak{r}(x_3x_2x_1) \subseteq \cdots,$$
$$\mathfrak{l}(Nx_1N) \subseteq \mathfrak{l}(Nx_2x_1N) \subseteq \cdots,$$

where $x_1, x_2, \cdots$ are in N (Shock, 1969).

If a nil subring S is nilpotent in R then there is an integer n such that $a_1 \cdots a_n = 0$ for all $a_1, \cdots, a_n$ in S and $b_1 \cdots b_{n-1} \neq 0$ for some appropriate $b_1, \cdots, b_{n-1}$ in S. The number n is called the **index** of the nilpotent subring S. Suppose all nil subrings are nilpotent in R. Is there an upper bound for the indices of the nilpotent subrings? The answer is no: Let T be the ring generated by $x_2, x_3, \cdots$ with the relations $(x_n)^n = 0$ and $(x_n)^{n-1} \neq 0$ for $n \geqq 2$, and furthermore $x_jx_k \neq x_kx_j$ for $j \neq k$. If an element y in T is nilpotent, then for some x_n each term of y is in $(x_n)^iT(x_n)^k$, where $i+k \geqq n$, or each term of y is in the subring generated by x_n. It follows that a nil subring is nilpotent. But $(x_{n+1})^n \neq 0$ implies the indices of the nilpotent subrings are unbounded. The infinite sum $x_2T + x_3T + \cdots$ is direct. So T is not a right finite dimensional ring. However, this is not the case in a right finite dimensional ring. *Let R denote a right finite dimensional ring. Let $Z(R)$ denote the right singular ideal of R. Then a nil subring S is nilpotent in R if and only if the subring $S \cap Z(R)$ is nilpotent in R. If $Z(R)$ is nilpotent, then a nil subring is not only nilpotent in R, but has index at most $k(\dim R+1)$, where k is the index of $Z(R)$* (Shock, 1969). We now prove the above results.

3. The prime radical. In 1943 R. Baer introduced the **lower nil radical** of a ring as a radical built from nilpotent rings [1]. We call an ideal M **prime** in R if $M \neq R$ and AB in M implies A or B is in M for any ideals A and B of R. N. McCoy first considered the intersection of all the prime ideals of a ring [14]. Then J. Levitzki showed that this intersection was the lower nil radical defined by R. Baer [13]. Hence the lower nil radical has become known as the **prime radical**. N. Jacobson gave the first elementwise characterization of the prime radical in terms of m-sequences [7, p. 195]. J.Lambek showed that the prime radical is the set of strongly nilpotent elements in R [10, p. 56]. An element x in R is **strongly nilpotent** if for each sequence $x_1, x_2, \cdots$ in R, where $x_1 = x$ and x_{k+1} is in x_kRx_k, we have $x_n = 0$ for some n. This is equivalent to saying that for each sequence $y_1, y_2, \cdots$ in R, where $y_1 = x$ and y_{k+1} is in $y_k \cdots y_1R$, we have $y_m \cdots y_1 = 0$ for some m. If x is strongly nilpotent, then the element xz is nilpotent for any z in R. Thus xR is nil and the prime radical is a nil subring. Let P denote the prime radical of R. The prime radical of the factor ring R/P is the zero ideal.

LEMMA 3.1. *Let x and y be in R such that xy is nilpotent and $xyx \neq 0$. Then* $\mathfrak{r}(x) \subsetneqq \mathfrak{r}(xyx)$.

Proof. We have $(xy)^n = 0$ and $(xy)^{n-1} \neq 0$ for some n. If $x(yx)^{n-1} \neq 0$, then

$xyx \neq 0$ implies $(yx)^{n-1}$ is in $\mathfrak{r}(xyx)-\mathfrak{r}(x)$. If $x(yx)^{n-1}=0$, then $xyx(yx)^{n-2}=0$ and $x(yx)^{n-2}\neq 0$, otherwise $(xy)^{n-1}=0$. Hence $(yx)^{n-2}$ is in $\mathfrak{r}(xyx)-\mathfrak{r}(x)$. In either case $\mathfrak{r}(x) \subsetneqq \mathfrak{r}(xyx)$.

PROPOSITION 3.2 (Shock [16]). *The prime radical of a ring R is the set of elements x in R such that xR is nil and the increasing sequence $\mathfrak{r}(x_1)\subseteq\mathfrak{r}(x_2x_1)\subseteq\mathfrak{r}(x_3x_2x_1)\subseteq \cdots$ becomes constant, where $x=x_1$ and x_{k+1} is in $x_k \cdots x_1R$.*

Proof. An element x in the prime radical is strongly nilpotent. Thus xR is nil and x has the desired sequential property. Conversely, let x and $x_1, x_2, \cdots$ be defined as in the hypothesis. It suffices to find some n such that $x_n \cdots x_1=0$, for then x would be strongly nilpotent. Suppose that $x_k \cdots x_1\neq 0$ for all $k\geqq 1$. Lemma 3.1 implies $\mathfrak{r}(x_1) \subsetneqq \mathfrak{r}(x_2x_1) \subsetneqq \cdots$ since $x_{k+1}x_k \cdots x_1=(x_k \cdots x_1)\cdot p(x_k \cdots x_1)$ for some p in R. This is a contradiction. Thus $x_n \cdots x_1=0$ for some n. This completes the proof.

Let $S_1\supseteq S_2\supseteq \cdots$ denote a decreasing sequence of subsets of R. We say that this sequence has a left **constant annihilator** if $\mathfrak{l}(S_1)\subseteq\mathfrak{l}(S_2)\subseteq \cdots$ becomes constant. It has a right constant annihilator if $\mathfrak{r}(S_1)\subseteq\mathfrak{r}(S_2)\subseteq \cdots$ becomes constant. We speak of a subring T of R as having a right constant annihilator if its sequence of powers $T\supseteq T^2\supseteq \cdots$ has a right constant annihilator. A nilpotent subring always has a right constant annihilator.

LEMMA 3.3. *Let P denote the prime radical of R. Let K be a subset in R where $(0)\subseteq K$. If x is in P and $Px\not\subseteq K$, then $yRyx\subseteq K$ and $yx\in P-K$ for some y in P.*

Proof. Pick b in P such that $bx\in P-K$. If $bRbx\not\subseteq K$, then pick b_1 in bRb such that $b_1x\in P-K$. If $b_1Rb_1x\not\subseteq K$, then pick b_2 in b_1Rb_1 such that $b_2x\in P-K$. If $b_2Rb_2x \not\subseteq K$, then repeat the procedure. This process can not continue since b is strongly nilpotent and $(0)\subseteq K$. Hence there is some y in P such that $yRyx\subseteq K$ and $yx\in P-K$.

THEOREM 3.4 (Shock [16]). *The prime radical P of a ring R is nilpotent if and only if P has a right constant annihilator and the sequence of principal ideals $Rx_1R\supseteq Rx_2x_1R\supseteq Rx_3x_2x_1R\supseteq \cdots$ has a left constant annihilator, where $x_1,x_2, \cdots$ are in P.*

Proof. Assume that P is nilpotent with index n. Then $P^n=(0)$ and P has a right constant annihilator. Furthermore, a sequence $Rx_1R\supseteq Rx_2x_1R\supseteq \cdots$, where x_i is in P has a left constant annihilator since $x_n \cdots x_1=0$. For the converse let $K=\mathfrak{r}(P^t)=\mathfrak{r}(P^{t+1})= \cdots$ for some t. Assume $P\not\subseteq K$ and let $x\in P-K$. If $Px\subseteq K$, then $P^tPx=(0)$ and x is in K, a contradiction. Thus $Px\not\subseteq K$. By Lemma 3.3 we have $yx\in P-K$ and $yRyx\subseteq K$ for some y in P. We recursively define a sequence where $y_1=yx$. Thus, $y_1\in P-K$ and $y_1Ry_1=y(xR)y_1\subseteq yRy_1\subseteq K$. Assume there is a finite sequence $y_1, \cdots, y_n$ such that $y_n \cdots y_1\in P-K$ and $y_hRy_h \cdots y_1\subseteq K$ for $1\leqq h\leqq n$. Let $b=y_n \cdots y_1$ and as before $Pb\not\subseteq K$. By Lemma 3.3 pick y_{n+1} in P such that $y_{n+1}Ry_{n+1}b\subseteq K$ and $y_{n+1}b\in P-K$. We conclude that there is a sequence $y_1, y_2, \cdots$ such that $y_k \cdots y_1\in P-K$ and $y_kRy_k \cdots y_1\subseteq K$ for all $k\geqq 1$. Let $h=t+2$, where $P^tK=(0)$. Define $x_1=y_h \cdots y_1$ and

$x_k = y_{hk} \cdots y_{h(k-1)+1}$ for $k \geqq 2$. We verify $x_k \cdots x_1 \in P-K$ and $x_kRx_k \cdots x_1 \subseteq P^tK=(0)$. The sequence $Rx_1R \supseteq Rx_2x_1R \supseteq \cdots$ does not have a left constant annihilator because $x_{k+1}Rx_{k+1} \cdots x_1R=0$ and $x_{k+1}x_k \cdots x_1 \neq 0$ for $k \geqq 1$. Thus $x_{k+1} \in \mathfrak{l}(Rx_{k+1} \cdots x_1R) - \mathfrak{l}(Rx_k \cdots x_1R)$. This contradicts the hypothesis. Thus $P \subseteq K$ and $P^{t+1}=(0)$ and P is nilpotent.

We combine our two results. This gives an annihilating condition for a nil ring to be nilpotent.

COROLLARY 3.5 (Shock [16]). *Let N denote a nil ring. Then N is nilpotent if and only if N has a right constant annihilator and the following sequences become constant*:

$$\mathfrak{l}(Nx_1N) \subseteq \mathfrak{l}(Nx_2x_1N) \subseteq \mathfrak{l}(Nx_3x_2x_1N) \subseteq \cdots \quad \textit{and} \quad \mathfrak{r}(x_1) \subseteq \mathfrak{r}(x_2x_1) \subseteq \cdots,$$

where $x_1, x_2, \cdots$ are in N.

Proof. If N is nilpotent then the implication is clear. For the converse let x be in N. Then $\mathfrak{r}(x_1) \subseteq \mathfrak{r}(x_2x_1) \subseteq \cdots$ becomes constant, where $x_1=x$ and x_{k+1} is in $x_k \cdots x_1N$. By Proposition 3.2 the prime radical of N is N. Theorem 3.4 implies N is nilpotent. This completes the proof.

We single out some useful facts about annihilators. Their proofs hinge only on definitions. For subsets A and B in R the relation $A \subseteq B$ implies $\mathfrak{r}(A) \supseteq \mathfrak{r}(B)$. Furthermore $\mathfrak{r}(\mathfrak{l}[\mathfrak{r}(A)]) = \mathfrak{r}(A)$. Therefore a strictly increasing sequence of right annihilators $\mathfrak{r}(S_1) \subsetneqq \mathfrak{r}(S_2) \subsetneqq \cdots$ implies a strictly decreasing sequence of left annihilators $\mathfrak{l}(\mathfrak{r}(S_1)) \supsetneqq \mathfrak{l}(\mathfrak{r}(S_2)) \supsetneqq \cdots$ and conversely. The minimum condition on left annihilators forces the maximum condition on right annihilators and conversely. The maximum condition on right annihilators is equivalent to saying that a decreasing sequence of sets has a right constant annihilator in R. Consequently:

If R satisfies the maximum condition on right annihilators, then a subring also enjoys this property.

Similarly, the result holds for left annihilators.

COROLLARY 3.6 (Herstein and Small [5]). *If R satisfies the maximum condition on right annihilators and on left annihilators, then a nil subring is nilpotent.*

Proof. A nil subring N has the maximum condition on right annihilators and on left annihilators. Corollary 3.5 implies N is nilpotent.

Recall a right ideal is essential if it has a nonzero intersection with each nonzero right ideal. If K and L are essential right ideals in R, then $K \cap H \neq (0)$ for any nonzero right ideal H in R. Also $L \cap (K \cap H) = (L \cap K) \cap H \neq 0$ since L is essential. Hence $K \cap L$ is essential. It follows by finite induction that the finite intersection of essential right ideals is an essential right ideal. Let $Z(R) = \{x : \mathfrak{r}(x)$ is essential$\}$. We now show that $Z(R)$, the right singular ideal of R, is indeed an ideal. If x and y are in $Z(R)$, then $\mathfrak{r}(x) \cap \mathfrak{r}(y)$ is essential. Hence $\mathfrak{r}(x+y) \supseteq \mathfrak{r}(x) \cap \mathfrak{r}(y)$ forces $x+y$ in $Z(R)$. Let $z \in Z(R)$ and $p \in R$. Then $\mathfrak{r}(pz) \supseteq \mathfrak{r}(z)$ implies $pz \in Z(R)$. Let H be a nonzero right ideal. If the right ideal H', where $H' = \{ph : h \in H\}$, is nonzero, then $\mathfrak{r}(z) \cap H' \neq (0)$ implies $\mathfrak{r}(zp) \cap H \neq 0$. Hence

$\mathfrak{r}(zp)$ is essential and $zp \in Z(R)$. This completes the proof.

COROLLARY 3.7 (Mewborn and Winton [15]). *Let R satisfy the maximum condition on right annihilators. Then the right singular ideal is nilpotent in R.*

Proof. Let $Z(R)$ denote the right singular ideal. The hypothesis implies $Z(R)$ has a right constant annihilator. Let $x_1, x_2, \cdots$ be in $Z(R)$. Suppose $x_n \cdots x_1 \neq 0$ for all n. Recall $\mathfrak{r}(x_{n+1})$ is essential and has a nonzero intersection with $x_n \cdots x_1 R$. Therefore $\mathfrak{r}(x_1) \subsetneqq \mathfrak{r}(x_2 x_1) \subsetneqq \cdots$, and this contradicts the hypothesis. Hence $x_n \cdots x_1 = 0$ for some n. This proves $Z(R)$ is nil and $Z(R)$ satisfies the annihilating conditions of Corollary 3.5. Thus $Z(R)$ is nilpotent.

COROLLARY 3.8 (Shock [16]). *If R has the minimum condition on ideals, then the prime radical is nilpotent.*

Proof. Let P denote the prime radical. Consider the sequences $P \supseteq P^2 \supseteq P^3 \supseteq \cdots$ and $Rx_1R \supseteq Rx_2x_1R \supseteq \cdots$, where $x_1, x_2, \cdots$ are in P. The minimum condition forces the sequences to become constant. Theorem 3.4 implies P is nilpotent.

4. Finite dimensional rings. Throughout this section let $Z(R)$ denote the right singular ideal of a ring R.

LEMMA 4.1. *Suppose $pa \in R - Z(R)$, where a and p are in R. Let $x_1, \cdots, x_n$ be in $Z(R)$. Then $x_1 z = \cdots = x_n z = 0$ for some z in R. Furthermore, $paz \neq 0$ and $\mathfrak{r}(paz) = \mathfrak{r}(az)$.*

Proof. The right annihilator $\mathfrak{r}(pa)$ is not essential. Thus $\mathfrak{r}(pa) \cap wR = (0)$ for some $0 \neq w$ in R. The finite intersection of essential right ideals is essential. For $1 \leqq i \leqq n$ the intersection $\bigcap \mathfrak{r}(x_i)$ is essential and meets wR. Hence for some $z \in wR - (0)$ we have $x_1 z = \cdots = x_n z = 0$ and $\mathfrak{r}(paz) = \mathfrak{r}(az)$. This completes the proof.

In the above lemma the case $x_1 = \cdots = x_n = 0$ is of interest. Then $pa \in R - Z(R)$ implies $paz \neq 0$ and $\mathfrak{r}(paz) = \mathfrak{r}(az)$ for some z in R. We use this fact below.

PROPOSITION 4.2 (Shock [17]). *Let $a_1, \cdots, a_n$ be in $R - Z(R)$ and let S_i contain $a_i, \cdots, a_n$ for $1 \leqq i \leqq n$. Assume that there are $p_1, \cdots, p_{n-1}$ in R such that $p_i a_i \in R - Z(R)$ and $p_i S_{i+1} \subseteq Z(R)$ for $1 \leqq i \leqq n-1$. [In the factor ring $\overline{R} = R/Z(R)$ this means $\mathfrak{l}(\overline{S}_1) \subsetneqq \cdots \subsetneqq \mathfrak{l}(\overline{S}_n)$ and $\mathfrak{l}(\overline{S}_n) \supsetneqq \cdots \supsetneqq \mathfrak{l}(\overline{S}_1)$.] Then $\dim R \geqq n$.*

Proof. First $p_1 a_1 \in R - Z(R)$ implies $p_1 a_1 z_1 \neq 0$ and $\mathfrak{r}(p_1 a_1 z_1) = \mathfrak{r}(a_1 z_1)$ for some z_1 in R. For $2 \leqq j \leqq n-1$ observe that $p_j a_j \in R - Z(R)$ and $p_1 a_j, \cdots, p_{j-1} a_j$ are in $Z(R)$. These conditions satisfy Lemma 4.1. Hence $p_1 a_j z_j = \cdots p_{j-1} a_j z_j = 0$ and $\mathfrak{r}(p_j a_j z_j) = \mathfrak{r}(a_j z_j)$; also $p_j a_j z_j \neq 0$ for some z_j in R. Also, $a_n \in R - Z(R)$ implies $\mathfrak{r}(a_n) \cap wR = 0$ for some $w \in R - (0)$. Since $p_i a_n \in Z(R)$ for $1 \leqq i \leqq n-1$, the finite intersection $\bigcap \mathfrak{r}(p_i a_n)$ is essential and meets wR. Therefore each $p_i a_n z_n = 0$ and $a_n z_n \neq 0$ for some z_n in wR. We now claim the sum $a_1 z_1 R + \cdots + a_n z_n R$ is direct. Suppose $0 \neq a_i z_i k_i = \sum a_j z_j k_j$ where $1 \leqq i < j \leqq n$ and k_i and k_j's are in R.

Then multiply on the left by p_i to get $0 \neq p_i a_i z_i k_i = 0$, because $p_i a_j z_j = 0$ and $\mathfrak{r}(p_i a_i z_i) = \mathfrak{r}(a_i z_i)$. This is a contradiction. The sum is direct and dim $R \geqq n$.

COROLLARY 4.3. *Let R be a right finite dimensional ring. Then the factor ring $R/Z(R)$ satisfies the maximum condition on both right and left annihilators.*

Proof. Recall that the minimum condition on left annihilators implies the maximum condition on right annihilators. The result follows from Proposition 4.2.

Corollary 4.3 is also stated in [3].

THEOREM 4.4 (Shock [17]). *Let R be a right finite dimensional ring. Then in the factor ring $R/Z(R)$ a nil subring is nilpotent and has index $\leqq$ dim $R+1$. A nil subring N is nilpotent in R if and only if the subring $N \cap Z(R)$ is nilpotent in R. Furthermore, if $Z(R)$ is nilpotent, then a nil subring is nilpotent in R and has index $\leqq k$(dim $R+1$) where $k =$ index of $Z(R)$.*

Proof. Corollary 4.3 followed by Corollary 3.6 imply that a nil subring $\overline{N}$ in $\overline{R} = R/Z(R)$ is nilpotent. Let t denote the index of $\overline{N}$. If $t > n+1$, where dim $R = n$, then $\mathfrak{r}(N) \subsetneqq \cdots \subsetneqq \mathfrak{r}(N^{t-1})$. By Proposition 4.2, dim $R \geqq t-1 > n$. This is a contradiction and thus $t \leqq n+1$. The proof now follows directly from the definition of a factor ring and a nilpotent subring.

COROLLARY 4.5 (Lanski [11]). *A nil subring is nilpotent in a right Goldie ring.*

Proof. Corollary 3.7 implies the right singular ideal is nilpotent in a right Goldie ring. Thus a nil subring is nilpotent by Theorem 4.4.

COROLLARY 4.6 (Levitzki [12]). *A nil right ideal is nilpotent in a right Noetherian ring.*

Proof. A right Noetherian ring is a right Goldie ring. The result follows from Corollary 4.5.

COROLLARY 4.7. *A nil subring is nilpotent in a right Artinian ring.*

Proof. Let R be a right Artinian ring. Let P denote the prime radical of R. First suppose $P = (0)$. If $x \in Z(R) - (0)$, then $xpx \neq 0$ for some p in R; otherwise $xRx = 0$ implies $x \in P$. Since $\mathfrak{r}(xp)$ is essential and meets xR, we have $xpxy = 0$ and $xy \neq 0$ for some y in R. Therefore $xp \in \mathfrak{l}(xy) - \mathfrak{l}(x)$ and $\mathfrak{l}(x) \subsetneqq \mathfrak{l}(xy)$. Apply this argument to xy. Again $xyp'xy \neq 0$ for some p' in R. As before $xyp'xyz = 0$ and $xyz \neq 0$ for some z in R. Hence $\mathfrak{l}(x) \subsetneqq \mathfrak{l}(xy) \subsetneqq \mathfrak{l}(xyz)$. The minimum condition on right ideals implies the maximum condition on left annihilators. This process cannot continue and thus $Z(R) = (0)$. By Theorem 4.4 a nil subring is nilpotent. If $P \neq (0)$, then P is nilpotent by Corollary 3.8. We have already shown that in R/P a nil subring is nilpotent. Therefore a nil subring is nilpotent in R.

COROLLARY 4.8 (Hopkins [6]). *A nil right ideal or a nil left ideal is nilpotent in a right Artinian ring.*

Proof. This follows from Corollary 4.7.

References

1. R. Baer, Radical ideals, Amer. J. Math., 65 (1943) 537–68.
2. N. J. Divinsky, Rings and radicals, University of Toronto Press, 1965.
3. J. Fisher, Finite dimensional rings, Notices, Amer. Math. Society, 17 (June 1970) 70T-A92.
4. A. W. Goldie, Semiprime rings with the maximum condition, Proc. London Math. Soc., 10 (1960) 201–220.
5. I. N. Herstein and L. Small, Nil rings satisfying certain chain conditions, Canad. J. Math., 16 (1964) 771–776.
6. C. Hopkins, Rings with minimal conditions for left ideals, Ann. Math., 40 (1939) 712–730.
7. N. Jacobson, Structure of rings, Amer. Math. Soc. Colloq., 37 (1968).
8. A. V. Jategaonkar, Left Principal Ideal Rings, Math. Lecture 123, Springer-Verlag, Berlin, 1970.
9. R. E. Johnson, Extended centralizer of a ring over a module, Proc. Amer. Math. Soc., 2 (1951) 891–895.
10. J. Lambek, Lectures on Rings and Modules, Blaisdell, Waltham, Mass., 1966.
11. C. Lanski, Nil subrings of Goldie rings are nilpotent, Canad. J. Math., 21 (1969) 904–907.
12. J. Levitzki, On multiplicative systems, Compositio Math., 8 (1960 (1950)) 76–80.
13. ———, Prime ideals and the lower radical, Amer. J. Math., 73 (1951) 25–29.
14. N. McCoy, Prime ideals in general rings, Amer. J. Math., 71 (1949) 823–833.
15. A. C. Mewborn and C. N. Winton, Orders in self-injective semiperfect rings, J. Algebra, 13 (1969) 5–10.
16. R. C. Shock, A note on the prime radical, submitted.
17. ———, Nil subrings of finite dimensional rings, submitted.

CONTINUITY AND BAIRE FUNCTIONS

E. R. LORCH, Columbia University

I. To attempt to define the various branches of mathematics, such as algebra, topology, analysis, is a rather breathtaking idea, but one which is quickly abandoned. However, it is quite feasible and very rewarding to try to give the flavor of these various disciplines. And this flavor is well described by setting down the vocabulary which is current in each of these subjects. Thus, surely, the words continuity, open, closed, metric, and compact, evoke the taste of topology. The simplest topological spaces are metric and if these spaces happen to be compact we may give a sigh of satisfaction. And the study of the properties of continuous functions defined on such spaces certainly belongs to the heart of the subject. Furthermore, the notion of continuity rests upon the concept of open set (the preimage of every open set is open). Thus in topology

Edgar Raymond Lorch wrote his Columbia Ph.D. thesis under J. F. Ritt. He spent a year at Harvard as an NRC Fellow and a year in Hungary as a Cutting Fellow before joining the Columbia Univ. staff (where he is now the Math. Dept. Chairman). He spent a Fulbright year at the Univ. of Rome and another at the Coll. de France; also he was a visiting professor at Carnegie Tech., Stanford, and METU, Ankara. His main research interests are functional analysis and set topology, and he is the author of *Spectral Theory* (Oxford Univ. Press 1962). *Editor*.

one studies properties of open sets. Closed sets are the complements of open sets and naturally come into play also.

To describe the flavor of analysis is considerably more difficult because of the variety of undertakings which come under this heading. The fact is that analysis is quite an octopus. But most would agree that integration belongs near the core of analysis. Also, characteristic of the subject is the predominant role played by denumerable operations: denumerable sums (infinite series); various limiting operations such as differentiation, which may be carried out by a countable process; denumerable unions and intersections of sets; to mention just a very few. Wherever one turns in analysis one is confronted with the cardinal $\aleph_0$ and with its extraordinary properties.

If one is allowed to proceed with these oversimplifying observations, one is carried along to the following considerations. Suppose one starts with a topological space $\boldsymbol{E}$. In this space one is presented with a particular class of sets, the open sets. As is customary one denotes an open set with the letter G. Closed sets are denoted with F. Since we have a topology, we may consider continuous real valued functions f—f: $\boldsymbol{E}\to\boldsymbol{R}$. If an integration theory is to be erected, it is reasonable to desire that many continuous functions be integrable. Now an integration process automatically involves a theory of measure. Thus we are confronted with a measure defined on an algebra of sets which (in the simple cases) includes open sets and closed sets. This algebra of sets necessarily includes the denumerable union and denumerable intersection of its members, that is, it is a σ-algebra. Thus if the sets G_n, $n=1, 2, \cdots$, are open, then $G_\delta=\bigcap_1^\infty G_n$ belongs to the algebra of sets on which the integration procedure rests. Similarly if the sets F_n, $n=1, 2, \cdots$, are closed, then $F_\sigma=\bigcup_1^\infty F_n$ belongs to this algebra of sets. The subscripts δ and σ stand for intersection (*Durchschnitt* in German) and union (*Summe* in German). Note that it is not necessary to consider such symbols as G_σ and F_δ since the union of any collection of open sets is open, not merely the union of a denumerable collection; and similarly for F_δ. Thus the algebra of sets on which the measure is defined contains not merely sets F and G but also sets F_σ and G_δ. Continuing in this way, one sees that it contains sets of the form $F_{\sigma\delta}$, $G_{\delta\sigma}$, $F_{\sigma\delta\sigma}$, $G_{\delta\sigma\delta}$, and so on. Note once more that any $F_{\sigma\sigma}$ is an F_σ and any $G_{\delta\delta}$ is a G_δ.

One may capsulate the preceding discussion. If one considers a topology on a space $\boldsymbol{E}$, one is concerned with a reasonably restricted class of sets, the open sets G and also the closed sets F. (Once in a while, a nonelementary theorem is phrased in terms of G_δ or F_σ, but almost never anything beyond.) These *generate* by the operations of denumerable union and intersection a σ-algebra of sets containing in addition to the sets F and G the sets F_σ, G_δ, $F_{\sigma\delta}$, $G_{\delta\sigma}$, and so on. This σ-algebra is usually enormously larger than the initial topology. The sets of the σ-algebra are called Borel sets. There is also defined the notion of Baire set. The notion of Baire set is a bit delicate. The family of Baire sets is the smallest σ-algebra which contains all compact, hence closed sets, which are simultaneously G_δ sets. It turns out that every Baire set is a Borel set but in general

there are Borel sets which are not Baire sets. However, for the special class of spaces which will be investigated below, the notion of Borel set coincides with that of Baire set. Whereas topology is interested in interrelations of the open and closed sets, analysis, more precisely the integration theory based on the topology, operates in a much larger domain, the σ-algebra generated by the topology. One proceeds from topology to analysis by crossing a bridge, and the symbol of that bridge is the word "generate." For the most part in the past, this trip over the bridge has been made blindfolded and in one direction only. We wish to suggest in what follows that the reverse direction should be attempted. In fact, we shall be led to study the class of topologies generating the same family of Baire sets. The passage from topology to σ-algebra is usually carried out blithely and at high speed. We wish to step in and try to investigate some of the relations involved.

It is possible to approach the subject from a slightly different point of view: that of Baire functions. Suppose $\{f_n : n=1, 2, \cdots\}$, represents a sequence of continuous real valued functions. Suppose that for every point x in the space E, the sequence $\{f_n(x)\}$ of real numbers is convergent. Write $\varphi(x)$ for the limit of this sequence. Thus $f_n(x) \to \varphi(x)$ as $n \to \infty$. Nothing is said about the nature of the convergence; in particular, it is not assumed to be uniform. Thus there is defined a real valued function φ and we may write $f_n \to \varphi$. The function φ is in general not continuous. It is said to be a Baire function of class 1. Consider now functions φ_n, $n=1, 2, \cdots$ of class 1 with the property that for each x in E, the sequence of real numbers $\{\varphi_n(x)\}$ is convergent. Write $\varphi_n(x) \to \psi(x)$ as $n \to \infty$. One thus has $\varphi_n \to \psi$ and ψ is said to be a Baire function of class 2. In this way, one may define Baire functions of classes 1, 2, 3, $\cdots$. In fact, the numbers describing the classes are ordinal numbers and thus it is possible to define functions of class ω, where ω is the first nonfinite ordinal, as the pointwise limit of a sequence of functions each of which belongs to a finite class. From there one goes on to functions of class $\omega+1, \omega+2, \cdots, \omega^2, \omega^2+1, \cdots$ and quite a bit beyond. Because of certain properties of ordinal numbers, the process is guaranteed to give nothing new when one reaches the ordinal number Ω that is the smallest ordinal number whose cardinality is nondenumerable. To round out the picture, the continuous functions are said to be of class 0. Thus the Baire functions are those functions defined by a point-wise limiting process starting with the continuous functions.

In order to bring the discussion in sharper focus, it is time to refer to one of the most classic of all topological spaces, the closed unit interval $[0, 1]$. Thus let E be the set of x such that $0 \leqq x \leqq 1$. Note that the absolute value defines a distance in E: the distance from x_1 to x_2 is given by $|x_2 - x_1|$. Note also that this distance function gives E a metric topology and that this topology is compact. For this space (in fact, for any metric space) the concepts of Borel set and Baire set coincide. As is well known, the Riemann integration process may be applied to any real-valued function continuous on $[0, 1]$. However, since there are Baire functions which are not Riemann integrable, (for example, the famous function

ψ, such that $\psi(x)=1$ when x is rational, and $\psi(x)=0$ when x is irrational, is a Baire function of class 2 but it is obviously not Riemann integrable) the Riemann integral has been displaced by the Lebesgue integral for which all bounded Baire functions are integrable. The Baire functions form the smallest class of functions containing the continuous functions for which the Lebesgue-Stieltjes integration theory is meaningful. Thus on the space $[0, 1]$ and with respect to this common type of integration, the class of Baire functions is *the* class which holds the spotlight.

The relation between Baire sets and Baire functions is rather simple. We shall state it here. (1) The characteristic function of any Baire set is a Baire function. (2) The uniform closure of the algebra of functions generated by the characteristic functions of Baire sets is precisely the algebra of bounded Baire functions. Thus the two theories: Baire sets and Baire functions are coextensive. One may proceed at one level or the other. We shall deal both with Baire sets and with Baire functions. The proof of (1) is straightforward. The proof of (2) requires a few preliminaries which will not be set down here.

II. It is of interest to remember a few data of historical importance. The notion of σ-algebra of sets precedes by some years that of topology. The names associated with the subject are almost exclusively French and belong to the golden years around 1900 of Parisian mathematics: Émile Borel, René Baire, Henri Lebesgue. Borel's *thèse* was presented to the Faculté des Sciences in Paris in 1894. Many of the ideas there developed are to be found in his book [2] published in 1898. Here one finds the notion of Borel measurable set (p. 46). The sets are not only measurable but have a real positive measure assigned to them. Here one finds also the famous Borel covering theorem for a closed interval (p. 42). It is proved in the form: any denumerable cover has a finite subcover. In his *thèse* he gives another proof. Borel exercised a powerful influence on his successors, not merely by virtue of his path-breaking researches, but also by his emphasis on the necessity of constructive solutions. It is worth noting that Borel does not introduce his sets as we do now: first open sets, then G_δ, then $G_{\delta\sigma}$, etc. (or first closed sets, then F_σ, then $F_{\sigma\delta}$, etc.). He considers on the line first intervals then countable disjoint unions of these, then complements with respect to these, then the countable disjoint union of these complements, and so on. This procedure has the advantage of making it easy to define the measure of a set, but it obscures its rank in the hierarchy of Borel sets.

The famous covering theorem is justly called the Heine-Borel theorem since the "continuous induction argument" had already been used over twenty years earlier by E. Heine. This student of Weierstrass, in an exposition of the properties of continuous functions, used the now standard compactness argument in order to prove the uniform continuity of a continuous function f defined on $[a, b]$. That is, he showed how starting at $x=a$, one can for a given $\epsilon>0$ proceed towards b and reach it in a finite number of interval steps so that the oscillation

of the function on any subinterval is less than ϵ (*Journal für die Reine und Angewandte Mathematik*, 74 (1872), p. 188).

René Baire, in his *thèse* published in 1899 [1], proceeds by functions instead of sets. His classification of functions is essentially that given in section I of this paper; however, he prefers to phrase his statements in terms of series rather than in terms of sequences. A very famous theorem is proved concerning functions of the first class. Functions of the second class are considered and results concerning them given. It is in the *thèse* that one finds the definition of sets of the first and second categories (p. 65), and the theorem that an interval on the real axis is of the second category (a complete metric space is of the second category.) His famous zero-dimensional space was introduced in 1909 in a paper in the *Acta Mathematica*.

The 1905 paper of Lebesgue is a masterpiece [6] from every point of view. In the first place it is replete with results of the highest luster. Next, it is written in incredibly beautiful and transparent style: the theorems are enunciated, italicized, and numbered. The proofs are cut into digestible segments, the whole exposition being carried out with the greatest elegance. To add to all this, which is already much too much, Lebesgue has the extraordinary good fortune of making a serious mistake, a real blooper. At a given point, he considers projections of Borel sets (from n dimensional to $n-k$ dimensional space) and states that the image of a Borel set under projection is a Borel set. "*Je vais démontrer que si E est mesurable B, sa projection e l'est aussi. Cela est évident si E est un intervalle car e en est un aussi. Or tout ensemble mesurable B se déduit d'intervalles par l'application répétée des opérations* I *et* II′, *lesquelles se conservent en projection ; la proposition est établie.*" Here operation I is denumerable union, operation II′ is denumerable monotone intersection. Now, the easiest kind of examples show that II′ does not commute with projection. This error was in large part the source of Lusin and Suslin's inspiration and led to their founding of the theory of analytic sets. Lebesgue in his preface to Lusin's book refers to it [12]:

"*A la réflection, une Préface m'a semblé être le seul endroit où je pourrais avouer très haut ce que M. Lusin a soigneusement caché: l'origine de tous les problèmes dont il va s'agir ici est une grossière erreur de mon Mémoire sur les fonctions représentables analytiquement. Fructueuse erreur, que je fus bien inspiré de la commettre!*" At the end of his book, Lusin writes a detailed analysis of Lebesgue's paper. This analysis is certainly one of the most extraordinary ever written of a mathematical memoir. Every idea is taken up, analyzed, and its deep implications are outlined. The treatment is with the utmost admiration and respect. Lusin even gives credit to Lebesgue for having discovered *le procédé des cribles* and to have given the first example of an analytic set. Nowhere does Lusin refer to the error. For him, Lebesgue is the master and personal teacher who does not make mistakes.

A principal result of the paper [6] is the proof of the existence of Baire classes of all orders (p. 205). One finds also much use of the notion of set of the first and second category and an extension to nondenumerable coverings of

Borel's compactness theorem (p. 176). This same extension was given simultaneously by F. Riesz (*Comptes Rendus*, Paris, 140 (1905), 224–226).

The next development is in the extension of these notions to abstract topological spaces. The fundamental classic in the introduction of set theoretic topology is Hausdorff's *Grundzüge der Mengenlehre* [3]. The publication of this book (first edition in 1914) also seems to mark the initiation in book form of modern axiomatic mathematics. The axioms for metric spaces occur on page 211 and the neighborhood axioms for a Hausdorff space are on page 213. In making this assertion we put aside previously known categorical axiom systems such as those of Hilbert for euclidean geometry. Hausdorff considers, among many other subjects, the structure of Borel sets and Baire functions, and one finds in the book many of the properties of these established in a general context. However, it is in the second edition (1927) called *Mengenlehre* [4] that one finds the full development of complete separable metric spaces and the principal results of the theory of Borel and Suslin sets.

III. The unit interval $[0, 1]$ is, as we know, at the heart of analysis and also of topology. However, as indicated above, it is not necessary nor is it fruitful in what follows for us to restrict ourselves exclusively to this space. The development which follows is valid in any compact metric space. The letter E will be used to refer to a set of points and the letter τ will refer to a compact metric topology on E. The topological space will be written (E, τ). The family of continuous real valued functions on (E, τ) will be indicated by C_τ. The family of bounded Baire functions associated with (E, τ) will be denoted by I_τ.

The discussion proceeds according to the cardinality of the set E. There are three cases:

(1) E has finite cardinality;
(2) E has denumerable cardinality;
(3) E has the cardinality $\mathfrak{c}$ of the continuum.

The basis of this classification rests on a theorem of topology which asserts that if a complete separable metric space has cardinality more than $\aleph_0$, then it has the cardinality $\mathfrak{c}$ of the continuum ([5], p. 445). The reason for this is that any such space contains a subset homeomorphic to the Cantor set. In addition, a compact metric space is necessarily separable, and a separable metric space cannot have cardinality greater than $\mathfrak{c}$. Thus the alternative to (1) and (2) is (3).

In case (1) the topology (E, τ) is discrete; the family C_τ consists of all functions defined on E, and $C_\tau = I_\tau$. Case (1) is essentially trivial. In case (2), there are many possible topologies; however, I_τ consists of all bounded functions on E. Case (2) calls for at most limited attention. The only case of real interest is (3). It should be noted that one finds under (3) a very large number of spaces of vital interest to mathematicians including such traditional objects as spheres in n-dimensions, tori, and so on.

To focus the development of the discussion, let us pose a problem. We start

as usual with a compact metric topology τ and have before us the families of functions C_τ and I_τ. Note that $C_\tau \subset I_\tau$. Now, suppose one were to pick a function f at random from I_τ. Are there ways of determining whether f belongs to C_τ? In other words, give conditions (necessary, sufficient, or both) that a Baire function be continuous. It so happens that the problem is incorrectly framed for reasons which will be apparent in a moment. But it is interesting to meditate on the fact that when correctly formulated, the problem is still unsolved in general; and the special cases in which a solution exists require principal and deep results from advanced set theory (analytic set theory).

The reason for stating that the above problem is not clearly formulated is the following. Starting with the fixed set E, one considers all compact metric topologies on E. For each such topology τ one constructs the families C_τ and I_τ. Now the map $\tau \rightarrow C_\tau$ is one-to-one as a consequence of an easy theorem of topology. However, the map $C_\tau \rightarrow I_\tau$ and hence the map $\tau \rightarrow I_\tau$ is *not* one-to-one. There are many compact metric topologies leading to the same class of Baire functions. This leads us to refine our procedures. Consider two topologies τ and τ' (always metric compact). We shall say that τ is equivalent to τ', $\tau \sim \tau'$, in case $I_\tau = I_{\tau'}$. This is indeed an equivalence relationship in the class of all compact metric topologies on E. If $\tau \sim \tau'$, we say that the topologies are coherent. If we start with a topology τ and subsequently construct a τ' such that $\tau \sim \tau'$, we shall say that the new topology τ' is coherent.

It is now apparent how the problem stated above can be well formulated. We start with a certain topology τ_0 which then defines an entire class of coherent topologies—all having the same Baire functions. Choosing a Baire function f, it is asked whether f is continuous for *some* coherent τ. An immediate first thought is that if f is continuous for some compact topology τ, then the range of f, the set $\{f(x): x \in E\}$, must be a compact (closed and bounded) set of real numbers. In case f separates points of E (that is, $x \neq y$ implies $f(x) \neq f(y)$), it can be shown that this condition is sufficient. In the general case, the situation is obscure.

We introduce now a concept fundamental to further development. Let (E_1, τ_1) and (E_2, τ_2) be two topological spaces. The reader will recall the notion of homeomorphism between these two. A homeomorphism is a one-to-one map Φ between E_1 and E_2 which carries τ_1 open sets of E_1 into τ_2 open sets of E_2 and whose inverse map Φ^{-1} carries τ_2 open sets of E_2 into τ_1 open sets of E_1. Two spaces which are homeomorphic are topologically the same: any topological property of one is possessed by the other. A *Baire isomorphism* between the two spaces is a one-to-one map Ψ between E_1 and E_2 which carries τ_1-Baire sets of E_1 into τ_2-Baire sets of E_2 and whose inverse map Ψ^{-1} carries τ_2-Baire sets of E_2 into τ_1-Baire sets of E_1. Two spaces which are Baire isomorphic have, in a sense, the same Baire structure, and there is a class of properties which they share in common. There is a theorem of topology which asserts that any two complete separable metric spaces with the same cardinality are Baire isomorphic ([5], p. 451). This theorem whose proof requires considerable labor has deep meaning

for the subsequent development.

Let E be a set of cardinality $\mathfrak{c}$ and let τ_0 be some fixed compact metric topology on E. A favorite choice for the pair (E, τ_0) is the closed unit interval $[0, 1]$ with the usual topology. Let T represent the family of all compact metric topologies τ coherent to τ_0, that is, $\mathrm{T}=\{\tau: I_\tau=I_{\tau_0}\}$. Let (E_1, τ_1) be any compact metric space of cardinality $\mathfrak{c}$—for example a spherical ball in 3-dimensional space. Then there exists a topology $\tau\in\mathrm{T}$ such that (E_1, τ_1) is homeomorphic to (E, τ). Thus T contains a representative of every compact metric space of cardinality $\mathfrak{c}$. The proof of this will be set down: let Ψ be any Baire isomorphism of (E_1, τ_1) and (E, τ_0); thus

$$\Psi: (E_1, \tau_1) \to (E, \tau_0). \tag{1}$$

Let τ be the unique topology defined on E such that

$$\Psi: (E_1, \tau_1) \to (E, \tau) \tag{2}$$

is a homeomorphism. This topology is easy to construct; the open sets of τ are the images by Ψ of the open sets of τ_1. Then τ is coherent to τ_0. This can be seen as follows. Let M be any τ-Baire set. Then $M_1=\Psi^{-1}(M)$ is a τ_1-Baire set in E_1 by the homeomorphism given in (2); next, $\Psi(M_1)$ is a τ_0-Baire set, since the map in (1) is a Baire isomorphism. But $\Psi(M_1)=\Psi(\Psi^{-1}(M))=M$. Thus any τ-Baire set M is also a τ_0-Baire set. The argument in the reverse direction is similar.

We therefore see that T includes among its elements every compact metric topology. In a certain sense therefore, T is a central object of analysis. In a moment we shall consider the question of giving T an adequate structure which allows one to initiate a fruitful study of it. The adequate structure will be a topology.

The essence of the introduction of T was the fact, stated above, that many distinct topologies have the same set of Baire functions. We wish to give life to this fact by adducing examples. First of all let E be a denumerable set. Then no matter what the topology τ is on E, I_τ consists of all bounded functions on E. Note that there are many metric compact topologies available; for example, topologies with precisely one, two, three, $\cdots$ limiting points. Going to the case of cardinality $\mathfrak{c}$, let (E, τ_0) be the closed unit interval $\{x: 0\leqq x\leqq 1\}$ with the usual topology. Now remove the point $x=1$ from $[0, 1]$ and put it on the side giving a half open interval $[0, 1)$ plus a discrete point. Next, make a loop out of $[0, 1)$ by attaching the two ends together. This gives a compact metric space (E, τ_1) consisting of a circle of circumference 1 plus a discrete point. The map which carries (E, τ_0) onto (E, τ_1) in the manner indicated above is a Baire isomorphism. Obviously, the two spaces are not homeomorphic.

It is possible to answer at this point some questions concerning the Baire functions and Baire sets associated with T. Let us agree, first of all, to write simply I instead of I_τ, since for any two topologies τ and τ' in T, one has $I_\tau=I_{\tau'}$. (Note also that there is a unique class of Baire sets associated with T.) We raise the question as to the extent to which the functions in I are continuous. A type

of answer has already been given to this question. We state now a result phrased in terms of approximation [8].

THEOREM A. *Let ϕ be a bounded Baire function and let ϵ be any positive number, $\epsilon>0$. Then there exists a coherent topology τ in* T *and a τ-continuous function f, where τ depends on ϵ, such that for all x in E,* $|\phi(x)-f(x)|<\epsilon$.

The theorem is interesting because of its paradoxical overtones. It states that Baire functions may be uniformly approximated by continuous functions. Now it is an old fact that the uniform limit of continuous functions is continuous. The difference between the classical and the new situation is, of course, that in the latter case as the approximation gets better, the topology changes. In the standard theorem τ is constant.

It is a classic result due to Lebesgue ([6], p. 205; also [4], p. 181), that if a compact metric space has the cardinality of the continuum, then there exist Baire sets of all orders α, where α is an ordinal number less than Ω, the first ordinal whose cardinality is nondenumerable. This means that given any α, there is a Baire set in E of order α and not of order less than α. Given this fact, let M be any Baire set associated with the class T. For every τ in T, the set M has a Baire order $\alpha(\tau)$ which depends on τ. A natural question to ask is: What is the range of the function $\alpha(\tau)$? What ordinal values may it assume and what values not? There are certain sets M for which this range is very limited. If, for example, M is a finite set, it is closed in any topology τ. Similarly, for denumerable sets and for sets with denumerable or finite complements, the range of $\alpha(\tau)$ is limited. Putting aside these exceptions, the range of $\alpha(\tau)$ is the entire set of ordinals less than Ω. We state this fact formally.

THEOREM B. *Let M be a Baire set such that both M and its complement have cardinality* c. *Let α be any ordinal number, $0\leqq\alpha<\Omega$. There exists a topology τ in* T *such that the Baire order of M with respect to τ is precisely α.*

This theorem suggests that, examined from the appropriate angle(choosing the appropriate τ), M has any preassigned degree of complexity. A reasonable presumption on our part is that the greater the value of α, the more complicated M is. In this connection, the reader is reminded of the fact that sets which are of type $F_{\sigma\delta\sigma}$ or $G_{\delta\sigma\delta}$ and not of lower type—with respect to any of the classic topologies—are very thinly scattered through the literature. In fact, looking for them is almost like hunting for unicorns. This fact gives one a feeling for the extraordinary degree of variation among the various coherent topologies.

IV. In order to penetrate further into this subject it is necessary to give an appropriate structure to T, the set of all coherent topologies. As mentioned earlier, this appropriate structure is itself a topology. This circumstance, that a collection of topologies is topologized, may seem a bit incestuous. However, a little thought will suggest its reasonableness. Two topologies τ and τ' could be said to be close to each other if they have many open sets in common; or if they have continuous functions in common. We shall choose the latter way. The

topology imposed on T will be called the *metatopology* to underline its cleavage in type from the host of topologies τ over which it rides herd.

The metatopology is defined by means of neighborhoods. Let τ_0 be any fixed point of T and let $f_1, \cdots, f_n$ be functions which are continuous in the topology τ_0. Then a typical neighborhood (of the base of neighborhoods) of τ_0 in the metatopology is defined by

$$\mathfrak{U}(\tau_0; f_1, \cdots, f_n) = \{\tau : \tau \in \mathrm{T} \text{ and } f_1, \cdots, f_n \text{ are } \tau\text{-continuous}\}. \tag{3}$$

Thus, this neighborhood of τ_0 consists of all topologies τ for which $f_1, \cdots, f_n$ are also continuous. It may be shown that the neighborhoods $\mathfrak{U}(\tau_0)$ defined in (3) are both open and closed and that the metatopology is separated. Thus T is totally disconnected.

The introduction of the metatopology raises a host of questions. We examine one or two. Suppose τ_0 is a compact metric topology such that there exist n τ-continuous functions $f_1, \cdots, f_n$ which separate points, that is, with the property that if the points x and y in E are distinct, then the n-tuples of real numbers $(f_1(x), \cdots, f_n(x))$ and $(f_1(y), \cdots, f_n(y))$ are distinct. In that case, the neighborhood of τ_0 defined by $\mathfrak{U}(\tau_0) = \mathfrak{U}(\tau_0; f_1, \cdots, f_n)$ contains the single point τ_0. This follows from elementary properties of compactness. Since $\mathfrak{U}(\tau_0)$ is an open set and since $\mathfrak{U}(\tau_0) = \{\tau_0\}$, the metatopology is discrete at the point τ_0. Thus the separation of points by a finite number of continuous functions implies discreteness. The converse proposition is not obvious but may be established [**10**].

It is easy to show that if for a given τ_0 there are τ_0-continuous functions $f_1, \cdots, f_n$ which separate points, then τ_0 is homeomorphic to a compact subset of n dimensional space $\boldsymbol{R}^n$. In fact, the map $x \to (f_1(x), \cdots, f_n(x))$ gives such a homeomorphism. The converse is easy to show (using coordinate functions). Putting these facts together with those of the preceding paragraph, we see that the metatopology is discrete at the point τ_0 if and only if τ_0 is homeomorphic to a compact subset of n-dimensional space. This throws interesting light on the problem of distinguishing topological objects by the process of describing nearby objects. The classical objects of analysis (spheres, etc., all of them n-dimensional) are of no value for this undertaking. For the only objects close to any given classical object is that object itself. What is needed here is a study of infinite dimensional objects, an undertaking which has not as yet borne fruit. As an example of a problem of central interest one can mention the following:

If τ is an infinite dimensional topology, does every neighborhood of τ contain finite dimensional topologies?

V. We have indicated in Section II the early history of the theory of Baire sets. There is a comparative lull in the field after 1905, probably due to the fact that most energies were devoted to extending the theory of the interval in $\boldsymbol{R}$ or $\boldsymbol{R}^n$ to general metric spaces. In 1917 there appeared two notes in the *Comptes Rendus*, written by the Russian mathematicians N. Lusin [**11**] and his pupil M. Suslin [**13**]. These authors had studied every aspect of Lebesgue's paper

and had found the *grossière erreur* described earlier. Undoubtedly they must have devoted considerable energy to proving Lebesgue's statement on projection; this would have closed the door on the great unknown and made the family of Borel sets the well fenced-in arena in which analysis could fight its battles without incursions from the outside. Substantial portions of Lebesgue's work could be validated by introducing new methods of proof. The projection theorem, however, could not be proved. In fact, the statement is false: There exist Borel sets in n-dimensional space whose projection to a lower dimensional space is not a Borel set ([**5**], p. 458). This fact is the initial and fundamental one in the construction of a new theory, the theory of analytic sets.

A fruitful definition of analytic set is: M is *analytic* if it is the continuous image of a Borel set ([**5**], p. 453). These sets and the sets formed by iterating the operations of complementation and continuous mapping are called projective sets. One finds here an awesome richness of phenomena. The projective sets are divided into classes in a manner somewhat similar to the Borel classes. One finds among classes of very low order, phenomena which go to the very root of the foundations of set theory. For example, the proposition that there exist sets of real numbers of the third projective class which are not Lebesgue measurable, is compatible with the Zermelo-Fraenkel set theory (all projective sets of order 0, that is all Borel sets, and all projective sets of orders 1 and 2 are Lebesgue measurable).

A last historical note: that most promising young mathematician, Suslin, who along with Lusin was the cofounder of the new theory, died at the very beginning of his scientific career. In 1919, at the age of 25, he succumbed to typhus. He had published one article of three pages [**13**]. He is also famous for a problem which appeared in his name in the *Fundamenta Mathematicae* 1 (1920), p. 223.

The principal result of the theory of analytic sets for our purposes is the following: *Let* $(\boldsymbol{E}, \tau)$ *and* $(\boldsymbol{F}, \sigma)$ *be two complete separable metric spaces. Let* Ψ *be a one-to-one map of* $\boldsymbol{E}$ *onto* $\boldsymbol{F}$. *Suppose that* Ψ *carries each Baire set of* $\boldsymbol{E}$ *into a Baire set of* $\boldsymbol{F}$. *Then* Ψ^{-1} *carries each Baire set of* $\boldsymbol{F}$ *into a Baire set of* $\boldsymbol{E}$. *Thus* Ψ *is a Baire isomorphism* ([**5**], p. 489). Since a compact metric space is separable and complete, the result applies to our situation. The metric character plays essentially in the hypothesis, since the result is not valid in arbitrary compact spaces. The above theorem is invoked in the proof of each of the theorems cited below.

Let us turn once more to one of the first problems mentioned in this paper. Let $f_1, \cdots, f_n$ be a finite number of bounded Baire functions. Consider the map Ξ from $\boldsymbol{E}$ to $\boldsymbol{R}^n$ given by

$$x \to (f_1(x), \cdots, f_n(x)), \qquad x \in \boldsymbol{E}. \tag{4}$$

This maps $\boldsymbol{E}$ onto a bounded subset of n-dimensional space. In case one considers a denumerable collection of bounded Baire functions $f_1, f_2, \cdots$, one considers instead that the map Ξ is from $\boldsymbol{E}$ into $\boldsymbol{R}^\omega$ and is defined by

$$x \to (f_1(x), f_2(x), \cdots), \qquad x \in \boldsymbol{E}. \tag{5}$$

Let $\Xi(\boldsymbol{E})$ denote the image of $\boldsymbol{E}$ under Ξ in the space $\boldsymbol{R}^n$ or $\boldsymbol{R}^\omega$ as the case may be. Note that by standard theorems of topology, the map Ξ, from the topological space $\boldsymbol{E}$ to the complete separable metric space $\boldsymbol{R}^n$ or $\boldsymbol{R}^\omega$, is continuous if and only if each of the functions f_i is continuous. Now if for a topology τ each of the functions f_i is continuous, then $\Xi(\boldsymbol{E})$ is a compact subset of $\boldsymbol{R}^n$ or $\boldsymbol{R}^\omega$ as the case may be. The reason for this is that Ξ is a continuous map and the continuous image of a compact space is compact. Thus the compactness of $\Xi(\boldsymbol{E})$ is necessary for $f_1, f_2, \cdots$ to be τ-continuous.

The question of sufficiency is posed in the following form: Let $f_1, f_2, \cdots$ be bounded Baire functions and let the map Ξ be defined as in (4) or (5). Suppose the set $\Xi(\boldsymbol{E})$ is compact. Does there exist a topology τ in T such that the functions $f_1, f_2, \cdots$ are τ-continuous? This question in its most general form has not been answered. However, the question can be answered positively in case the functions $f_1, f_2, \cdots$ separate the points of $\boldsymbol{E}$. This as yet unpublished result will be stated formally:

THEOREM C. *Let $f_1, f_2, \cdots$ be a finite or denumerable collection of bounded Baire functions in I, $m = 1, 2, \cdots$. Let the functions $f_1, f_2, \cdots$ separate the points of $\boldsymbol{E}$. Then a necessary and sufficient condition that there exist a compact metric topology τ in* T *such that the given functions f_m be τ-continuous is that the image of the map Ξ defined in* (4) *or* (5) *be compact.*

Theorem C may be stated in a different form. Any set of τ-continuous functions generates, by the formation of sums and products and the taking of uniform limits, a Banach algebra A of τ-continuous functions. The norm $\|f\|$ of a function $f \in A$ is defined by $\|f\| = \sup |f(x)|$, $x \in \boldsymbol{E}$. We shall assume that the algebra A contains the constant functions. In case this algebra separates the points of $\boldsymbol{E}$, then A is the algebra of all τ-continuous functions, that is, $A = C_\tau$. We remind ourselves that the algebra C_τ considered as a metric space is separable. That is, there exists a sequence $\{f_n\}$ which is dense in C_τ. Without going into further explanations we restate the previous result.

THEOREM C'. *Let A be a Banach algebra of bounded Baire functions in I containing the constants and with the supremum norm. Suppose that the functions of A separate the points of $\boldsymbol{E}$. Then there exists a compact metric topology $\tau \in$* T *such that $A = C_\tau$ if and only if there exists a denumerable set $\{f_n\}$ dense in A such that the image of $\boldsymbol{E}$ by the map Ξ given in* (5) *is compact.*

VI. An object of major attention in the study which has been undertaken so far is the family T of compact metric topologies which generate the same algebra of Baire functions. Of interest at least as great is the family of Baire automorphisms which are present at every step of the undertaking. These will now be considered. The Baire automorphisms are the bijective (one-to-one) maps of $\boldsymbol{E}$ onto $\boldsymbol{E}$ which, along with their inverses, carry Baire sets into Baire

sets. The letters $\mathfrak{g}$, $\mathfrak{h}$, $\mathfrak{k}$, and so on, will be used to denote them. In the first place, it is obvious that the identity map $\mathfrak{e}: E \rightarrow E$ which carries each element $x \in E$ into itself ($\mathfrak{e}x = x$) is a Baire automorphism. Next if $\mathfrak{g}$ is a Baire automorphism, its inverse $\mathfrak{g}^{-1}$ is also one. Finally, the product $\mathfrak{g} \cdot \mathfrak{h}$ of the two Baire automorphisms $\mathfrak{g}$ and $\mathfrak{h}$, defined by $(\mathfrak{g} \cdot \mathfrak{h})x = \mathfrak{g}(\mathfrak{h}x)$, is clearly a Baire automorphism. Thus these automorphisms form a group $\mathfrak{G}$. It is this group which will now be examined.

The group $\mathfrak{G}$ is very large. Such large algebraic objects usually cannot be studied successfully unless one attaches to them another structure. The usual structure to assign is a topological one; in the most successful cases, a topological group results. The question of assigning a topology to $\mathfrak{G}$ is not clear cut. Choices necessarily have to be made which may seem rather arbitrary. We shall introduce two types of topologies below. The "correctness" of the choice is then supported by the proof of a completeness theorem. The topologies we introduce on $\mathfrak{G}$ are closely related to the metatopology on T and it is this fact which has strongly influenced our choices.

Let us state at the outset that there will be two topologies introduced on $\mathfrak{G}$ for every compact metric topology τ in T. The topological spaces which result are denoted by ${}_\tau\mathfrak{G}$ and ${}^\tau\mathfrak{G}$. The topologies are defined by describing the neighborhoods of the identity $\mathfrak{e}$ of $\mathfrak{G}$. The neighborhood of an arbitrary element $\mathfrak{g}$ in $\mathfrak{G}$ is then given by performing a left group translation by $\mathfrak{g}$ on the various neighborhoods of the identity. Naturally, there are also right topologies $\mathfrak{G}_\tau$ and $\mathfrak{G}^\tau$ given by right group translations. (Group translation by $\mathfrak{g}$ means group multiplication by $\mathfrak{g}$.)

We start with a fixed τ in T. Let $f_1, \cdots, f_n$ be any τ-continuous functions. These functions may or may not distinguish points of E. Let x_0 be any point of E, and suppose that $f_1(x_0) = \alpha_1, \cdots, f_n(x_0) = \alpha_n$. Let us return to the map Ξ given in (4) which maps (E, τ) into R^n. Thus the point $(\alpha_1, \cdots, \alpha_n)$ in R^n is the image of $x_0 \in E$ by the continuous map Ξ. Consider the preimage of $(\alpha_1, \cdots, \alpha_n)$ under Ξ, that is, consider the set

$$M = \Xi^{-1}(\alpha_1, \cdots, \alpha_n). \tag{6}$$

The set will be called a *set of indeterminacy associated with the functions* $f_1, \cdots, f_n$. To obtain all sets of indeterminacy, one allows $(\alpha_1, \cdots, \alpha_n)$ to range over all points in the range of the map Ξ.

The set M defined in (6) depends on the n-tuple $(\alpha_1, \cdots, \alpha_n)$; thus we write $M = M(\alpha_1, \cdots, \alpha_n)$. Now let $\mathfrak{h}$ be a Baire automorphism in $\mathfrak{G}$ which transforms each $M(\alpha_1, \cdots, \alpha_n)$ into itself. The totality of all these Baire automorphisms $\mathfrak{h}$ is a subgroup $\mathfrak{H}$ of $\mathfrak{G}$. We write $\mathfrak{H} = \mathfrak{H}(\tau; f_1, \cdots, f_n)$. The subgroup character of $\mathfrak{H}$ is obvious. This subgroup $\mathfrak{H}$ is by definition a neighborhood of $\mathfrak{e}$ for the topology ${}_\tau\mathfrak{G}$. The set of all neighborhoods obtained by varying $f_1, \cdots, f_n$ in all possible ways constitutes the base of neighborhoods of $\mathfrak{e}$ for the topology ${}_\tau\mathfrak{G}$. It can be shown easily that the topology ${}_\tau\mathfrak{G}$ is discrete if and only if there exist τ-continuous functions $f_1, \cdots, f_n$ which separate points. Thus the interesting

topologies on $\mathfrak{G}$ correspond precisely to the interesting points τ in T.

The topology ${}_\tau\mathfrak{G}$ is a uniform topology; that is, it is associated with a uniform structure on $\mathfrak{G}$. This being the case, it is possible to raise certain questions concerning ${}_\tau\mathfrak{G}$ which are appropriate to uniform spaces. The most immediate one of these concerns the property of completeness. As is well known, in a metric space it is possible to ask the question: is it complete? What this means is: does every Cauchy sequence in the space converge to a point in the space? A similar question can be asked in any space endowed with a uniform topology. The only difference is that the notion of Cauchy sequence has to be replaced by something more powerful. This more powerful concept is that of generalized Cauchy sequence defined on directed sets. A uniform space is said to be complete providing that every generalized Cauchy sequence converges to a point in the space. It is possible to establish this property for the space ${}_\tau\mathfrak{G}$:

THEOREM D. *The uniform space* ${}_\tau\mathfrak{G}$ *is complete. Thus, if* $\{\mathfrak{g}_m\}$ *is any generalized Cauchy sequence, there exists a Baire automorphism* $\mathfrak{g}$ *in* $\mathfrak{G}$ *such that* $\{\mathfrak{g}_m\}$ *converges to* $\mathfrak{g}$.

The proof of this theorem is carried out in a series of steps which cannot be given here (see [7], pp. 142, 143). The entire apparatus developed up to this point enters into the discussion. In particular, the theorem on analytic sets which was enunciated in Section V is indispensable.

The topology ${}_\tau\mathfrak{G}$ is rather strong from certain points of view. This is due to the fact that a typical neighborhood of the identity e, $\mathfrak{H}(\tau, f_1, \cdots, f_n)$, consists of Baire automorphisms $\mathfrak{h}$ which transform each set M of indeterminacy in (6) into itself. Strong topologies have many open sets. In the most extreme case the topology is discrete, and this is usually not of interest since the closure operation is trivial.

We introduce below a uniform topology in $\mathfrak{G}$, denoted by ${}^\tau\mathfrak{G}$, which is weaker than ${}_\tau\mathfrak{G}$. Since ${}^\tau\mathfrak{G}$ is weaker, the neighborhoods of $\mathfrak{e}$ are fewer in number and hence there are more Cauchy sequences. Nevertheless, even though the number of these sequences is larger, the completeness of ${}^\tau\mathfrak{G}$ may also be established.

The typical neighborhoods of the identity $\mathfrak{e}$ for the topology ${}^\tau\mathfrak{G}$ are defined as follows. Let τ be fixed as usual. Let $f_1, \cdots, f_n$ be any τ-continuous functions. Consider all the sets of indeterminacy $M(\alpha_1, \cdots, \alpha_n)$ associated with $f_1, \cdots, f_n$. Consider now a Baire automorphism $\mathfrak{k}$ which *permutes* these sets. Thus, if $(\alpha_1, \cdots, \alpha_n)$ and $(\beta_1, \cdots, \beta_n)$ are two points in the range of the map defined by $x \rightarrow (f_1(x), \cdots, f_n(x))$, then if $\mathfrak{k}$ carries one point of $M(\alpha_1, \cdots, \alpha_n)$ into a point of $M(\beta_1, \cdots, \beta_n)$, it maps all of the former set onto the latter set. In particular, the two sets have the same cardinality. The totality of these Baire automorphisms $\mathfrak{k}$ is obviously a group which will be denoted by $\mathfrak{K}(\tau; f_1, \cdots, f_n)$. The family of such groups, obtained by varying $f_1, \cdots, f_n$ in all possible ways, gives a base of neighborhoods of the identity $\mathfrak{e}$. The neighborhoods of an arbitrary element $\mathfrak{g}$ in $\mathfrak{G}$ are obtained by left group multiplication by $\mathfrak{g}$ of the neighborhoods $\mathfrak{K}$ of the origin.

The topology ${}^{\tau}\mathfrak{G}$ thus obtained arises once more from a uniformity. It is not difficult to see that ${}_{\tau}\mathfrak{G}$ is stronger than ${}^{\tau}\mathfrak{G}$. Once more, it is possible to prove the completeness of the topology ${}^{\tau}\mathfrak{G}$. The completeness proof consists of four steps which are the same as those occurring in the proof of theorem D. However, one of the steps in the present situation requires much more delicate handling than that needed in the case of ${}_{\tau}\mathfrak{G}$. The degree to which the compactness and metrizability properties are involved and the intertwining character of the elements in the proofs suggest the "correctness" of the study of the topologies ${}_{\tau}\mathfrak{G}$ and ${}^{\tau}\mathfrak{G}$.

In view of the solitary character of researches in the highly unexplored areas here described, mathematico-esthetic criteria play an important role in helping to establish validity. In addition to these general criteria there is a very important guiding principle. At the heart of mathematical activity are the natural numbers N and the closed unit interval $[0, 1]$. Along with these go their denumerable powers, the Baire space N^{ω} (usually written N^{N}) and the Hilbert cube $[0, 1]^{\omega}$. Any undertaking which is aimed at revealing the properties of these fundamental entities is *ipso facto* well directed. This is the meeting ground for topology, analysis, and also logic. It is our feeling that this area will attract again, and with great fruitfulness, the attention it had in the first quarter of this century.

Text of an address delivered to the New York Metropolitan Section, MAA, April 18, 1970.

References

1. R. Baire, Sur les fonctions des variables réelles, Annali. Mat. III, 3 (1899) 1–123.
2. E. Borel, Leçons sur la théorie des fonctions, Paris (1898) 136 pp.
3. F. Hausdorff, Grundzüge der Mengenlehre, 1st ed., Leipzig, 1914, 476 pp.
4. ———, Mengenlehre, 2nd ed., Leipzig, 1927, 285 pp.
5. C. Kuratowski, Topology, vol. 1, Academic Press, New York, 1966, 560 pp.
6. H. Lebesgue, Sur les fonctions représentables analytiquement, J. Math., Ser. 16, 1(1905) 139–216.
7. E. R. Lorch, On compact metric spaces and the group of Baire equivalences, Studia Math., 31 (1968) 137–144.
8. ——— and Hing Tong, Continuity of Baire functions and order of Baire sets, J. Math. Mech., 16 (1967) 991–996.
9. ———, A completeness theorem on the group of Baire equivalences, J. Math. Mech., 19 (1969) 189–194.
10. ———, Caractérisation de certaines topologies compactes, L'Enseignement Math., 16 (1969) 153–157.
11. N. Lusin, Sur la classification de M. Baire, C. R. Acad. Sci. Paris, 164 (1917) 91–94.
12. ———, Leçons sur les ensembles analytiques et leurs applications, Paris, 1930, 326 pp.
13. M. Suslin, Sur une définition des ensembles mesurables B sans nombres transfinis, C.R. Acad. Sci. Paris, 164 (1917) 88–91.

THE WILLIAM LOWELL PUTNAM MATHEMATICAL COMPETITION

J. H. McKAY, Oakland University

The following results of the thirty-first William Lowell Putnam Mathematical Competition held on December 5, 1970, have been determined in accordance with the regulations governing the Competition. This competition is supported by the William Lowell Putnam Intercollegiate Memorial Fund left by Mrs. Putnam in memory of her husband and is held under the auspices of the Mathematical Association of America.

The first prize, five hundred dollars, is awarded to the Department of Mathematics of the **University of Chicago,** Chicago, Illinois. The members of the team were Robert Israel, Robert A. Oliver, and Robert Tax; to each of these a prize of one hundred dollars is awarded.

The second prize, four hundred dollars, is awarded to the Department of Mathematics of **Massachusetts Institute of Technology,** Cambridge, Massachusetts. The members of the team were David M. Christie, Don Coppersmith, and Steven Winker; to each of these a prize of seventy-five dollars is awarded.

The third prize, three hundred dollars, is awarded to the Department of Mathematics of the **University of Toronto,** Toronto, Ontario, Canada. The members of the team were Daniel Gautreau, Daryl Geller, and Joseph S. Repka; to each of these a prize of fifty dollars is awarded.

The fourth prize, two hundred dollars, is awarded to the Department of Mathematics of the **Illinois Institute of Technology,** Chicago, Illinois. The members of the team were George F. Cornelius, Zbigniew Friedorowicz, and Wayne F. Mroz; to each of these a prize of fifty dollars is awarded.

The fifth prize, one hundred dollars, is awarded to the Department of Mathematics of the **California Institute of Technology,** Pasadena, California. The members of the team were Leonidas Guibas, Andrew M. Odlyzko, and David J. Smith; to each of these a prize of fifty dollars is awarded.

The six persons ranking highest in the examination, named in alphabetical order, are **Jockum Aniansson,** Yale University; **Don Coppersmith,** Massachusetts Institute of Technology; **Jeffrey Lagarias,** Massachusetts Institute of Technology; **Robert A. Oliver,** University of Chicago; **Arthur Rubin,** Purdue University; and **Steven K. Winker,** Massachusetts Institute of Technology. Each of these has been designated as a Putnam fellow by the Mathematical Association of America and is awarded a prize of two hundred and fifty dollars.

The five persons ranking second highest in the examination, named in alphabetical order, are *Daryl Geller*, University of Toronto; *Zbigniew Friedorowicz*, Illinois Institute of Technology; *Dale H. Peterson*, Yale University; *Joseph S. Repka*, University of Toronto; and *Jonathan Rosenberg*, Harvard University. To each of these a prize of one hundred dollars is awarded.

The following teams, named in alphabetical order, won honorable mention: *University of California at Davis*, the members of the team were Dean Hickerson,

Peter Loomis, and James Howell; *Harvard University*, the members of the team were Avner Ash, Gerald Myerson, and Jonathan Rosenberg; *Princeton University*, the members of the team were Allen H. Back, James R. Paulson, and Steven Weintraub; *Reed College*, the members of the team were Joe P. Buhler, Thor Wenzel, and Dean Alvis; *Yale University*, the members of the team were Frederic B. Weissler, Jockum Aniansson, and Dale Peterson.

Honorable mention is given to the following twenty-nine individuals, named in alphabetical order: Stuart M. Ambler, *Harvard University;* Robert M. Anderson, *University of Toronto;* Timothy J. Augustine, *University of Notre Dame;* Allen H. Back, *Princeton University;* Kent M. Brothers, *University of Victoria;* Joe P. Buhler, *Reed College;* David M. Christie, *Massachusetts Institute of Technology;* George F. Cornelius, *Illinois Institute of Technology;* Paul F. Daniels, *Williams College;* Ken R. Davidson, *University of Waterloo;* Bruce E. Ferrero, *Cornell University;* David S. Fried, *University of Chicago;* Nadine P. Goldberg, *State University of New York at Buffalo;* Phillip P. Green, *Harvard University;* Dean R. Hickerson, *University of California at Davis;* Robert B. Israel, *University of Chicago;* Peter Loomis, *University of California at Davis;* John Mallet-Paret, *University of Alberta;* Lawrence D. Meisel, *Yale University;* Mark D. Meyerson, *University of Maryland;* Andrew M. Odlyzko, *California Institute of Technology;* Eric Rosenthal, *Yale University;* Jonathan Schonfeld, *Yale University;* James R. Spriggs, *Case Western Reserve University;* Jonathan Sussman, *San Diego State College;* Garrett S. Sylvester, *Princeton University;* Robert E. Tax, *University of Chicago;* Lawrence C. Washington, *Johns Hopkins University;* Paul J. Weiner, *Harvard University.*

The other individuals who ranked in the top one hundred, arranged by college, are: Alfred R. Weiss, *University of Alberta;* Peter J. Oliver and Alan J. Tausch, *Brown University;* Bill N. Celmaster, *University of British Columbia;* Max Marshall, *University of California at San Diego;* Paul F. Klembeck, *University of California at Los Angeles;* Darell J. Johnson, *University of California at Riverside;* Steven G. Krantz, *University of California at Santa Cruz;* William K. Delaney, Leonidas J. Guibas, Bruce A. Reznick, David J. Smith, and Michael F. Yoder, *California Institute of Technology;* Paul A. Carlson, *Carleton University;* Walter O. Augenstein, John A. MacBain, and Michael Somos, *Case Western Reserve University;* Stephen L. Millman, *Dartmouth College;* William H. Beckmann, *Davidson College;* Avner D. Ash, Ira M. Gessel, David Harbater, Joseph Harris, Peter A. Masters, Gerald I. Myerson, and John P. Robertson, *Harvard University;* Timothy J. Keller, *Harvey Mudd College;* Gabor Fencsik, *University of Illinois;* Wayne F. Mroz, *Illinois Institute of Technology;* Harold D. Taylor, and Max M. Wells, *University of Kansas;* Richard A. Arratia, Peter A. Gerritson, Steven F. McKay, Donald S. Raila, and Kenneth P. Rietz, *Massachusetts Institute of Technology;* John S. Pettengill, *University of Michigan;* Ben A. Murray, *Michigan State University;* Serge Hamelin, *University of Montreal;* Kenneth A. Brakke, *University of Nebraska;* Randall W. Mercer,

University of New Mexico; Dave Schmitz, *University of North Dakota;* Richard F. Poppen, *Pomona College;* Sheldon J. Axler, Peter Ochshorn, and James R. Paulson, *Princeton University;* David S. Jerison, *Purdue University;* Malcolm P. Hamilton, *Queen's University;* Kevin W. Kadell, *Sacramento State College;* Stewart C. Strait, *San Diego State College;* Daniel H. Leucking, *Southern Illinois University at Edwardsville;* William H. Rowan, *Stanford University;* William R. Franklin and Daniel A. Gautreau, *University of Toronto;* George S. Lueker, *Valparaiso University;* Chris J. Odgers, *University of Victoria;* Frank R. Allaire, *University of Waterloo;* Walter B. Rassbach, *Wesleyan University;* Kenneth P. Baclawski, *University of Wisconsin at Milwaukee.*

One thousand four hundred and forty-five students from two hundred and ninety-eight colleges and universities participated in the examination on December 5, 1970.

A listing of the top five hundred contestants may be obtained from the Director. The list, which includes addresses and expected dates of graduation, may be helpful to departments of mathematics in selecting graduate students.

The Questions Committee, consisting of Albert Wilansky (chairman), Warren Loud, and Murray S. Klamkin prepared the problems (listed below) for the competition.

PROBLEMS. PART A

A-1. Show that the power series for the function

$$e^{ax}\cos bx \qquad (a>0,\, b>0)$$

in powers of x has either no zero coefficients or infinitely many zero coefficients.

A-2. Consider the locus given by the real polynomial equation

$$Ax^2+Bxy+Cy^2+Dx^3+Ex^2y+Fxy^2+Gy^3=0,$$

where $B^2-4AC<0$. Prove that there is a positive number δ such that there are no points of the locus in the punctured disk

$$0<x^2+y^2<\delta^2.$$

A-3. Find the length of the longest sequence of equal nonzero digits in which an integral square can terminate (in base 10) and find the smallest square which terminates in such a sequence.

A-4. Given a sequence $\{x_n\}$, $n=1, 2, \cdots$, such that $\text{limit}_{n\to\infty}\{x_n-x_{n-2}\}=0$.
Prove that

$$\underset{n\to\infty}{\text{limit}}\frac{x_n-x_{n-1}}{n}=0.$$

A-5. Determine the radius of the largest circle which can lie on the ellipsoid

$$\frac{x^2}{a^2}+\frac{y^2}{b^2}+\frac{z^2}{c^2}=1 \qquad (a>b>c).$$

A-6. Three numbers are chosen independently at random, one from each of the three intervals $[0, L_i]$ $(i=1, 2, 3)$. If the distribution of each random number is uniform with respect to length in the interval it is chosen from, determine the expected value of the smallest of the three numbers chosen.

PART B

B-1. Evaluate

$$\lim_{n\to\infty} \frac{1}{n^4} \prod_{i=1}^{2n} (n^2 + i^2)^{1/n}.$$

B-2. The time-varying temperature of a certain body is given by a polynomial in the time of degree at most three. Show that the average temperature of the body between 9 A.M. and 3 P.M. can always be found by taking the average of the temperatures at two fixed times, which are independent of which polynomial occurs. Also, show that these two times are 10:16 A.M. and 1:44 P.M. to the nearest minute.

B-3. A closed subset S of R^2 lies in $a < x < b$. Show that its projection on the y-axis is closed.

B-4. An automobile starts from rest and ends at rest, traversing a distance of one mile in one minute, along a straight road. If a governor prevents the speed of the car from exceeding ninety miles per hour, show that at some time of the traverse the acceleration or deceleration of the car was at least 6.6 ft./sec.^{2}

B-5. Let u_n denote the "ramp" function

$$u_n(x) = \begin{cases} -n & \text{for } \quad x \leqq -n, \\ x & \text{for } -n < x \leqq n, \\ n & \text{for } \quad x > n, \end{cases}$$

and let F denote a real function of a real variable. Show that F is continuous if and only if $u_n \circ F$ is continuous for all n. (Note: $(u_n \circ F)(x) = u_n[F(x)]$.)

B-6. A quadrilateral which can be inscribed in a circle is said to be *inscribable* or *cyclic*. A quadrilateral which can be circumscribed to a circle is said to be *circumscribable*. Show that if a circumscribable quadrilateral of sides a, b, c, d has area $A = \sqrt{abcd}$, then it is also inscribable.

SOLUTIONS. PART A

The number in parentheses, immediately following the problem number, is the number of participants who received a score of 8, 9 or 10 (10 is the maximum possible) on the problem. In the case of A-1, A-2, B-1, and B-2, this applies to all 1445 participants. For the other problems, the count applies only to the 728 qualifiers.

A-1 (171). Note that $e^{ax} \cos bx$ is the real part of $e^{(a+ib)x}$. Thus the power series is

$$e^{ax} \cos bx = \sum_{n=0}^{\infty} \operatorname{Re}\{(a+ib)^n\} \frac{x^n}{n!} \cdot$$

In this form, it is easily seen that if x^n has a zero coefficient, then x^{kn} has a zero coefficient for every odd value of k.

A-2 (88). Let $(x, y) = (r \cos \theta, r \sin \theta)$, $r > 0$, be a point of the locus. Then

$$(1) \qquad r = \frac{|A \cos^2 \theta + B \sin \theta \cos \theta + C \sin^2 \theta|}{|D \cos^3 \theta + E \cos^2 \theta \sin \theta + F \cos \theta \sin^2 \theta + G \sin^3 \theta|} \cdot$$

The denominator of (1) is less than or equal to $|D| + |E| + |F| + |G|$, whereas the numerator has a positive minimum

$$N = \frac{|A+C| - \sqrt{(A-C)^2+B^2}}{2},$$

since $B^2<4AC$. Therefore

$$r \geqq \frac{N}{|D|+|E|+|F|+|G|} = \delta$$

and there are no points of the locus within $0<r<\delta$.

Alternate Solution: Set $H(x, y)$ equal to the polynomial on the left hand side of the given equation. The standard theory for maxima or minima of functions of two variables can be used together with the condition $B^2<4AC$ to show that $H(x, y)$ has a local maximum or a local minimum at (0, 0).

A-3 (85). If x is an integer then $x^2\equiv 0$, 1, 4, 6 or 9 (mod 10). The case $x^2\equiv 0$ (mod 10) is eliminated by the statement of the problem. If $x^2\equiv 11$, 55 or 99 (mod 100), then $x^2\equiv 3$ (mod 4) which is impossible. Similarly, $x^2\equiv 66$ (mod 100) implies $x^2\equiv 2$ (mod 4) which is also impossible. Therefore $x^2\equiv 44$ (mod 100). If $x^2\equiv 4444$ (mod 10,000), then $x^2\equiv 12$ (mod 16), but a simple check shows that this is impossible. Finally note that $(38)^2=1444$.

A-4 (110). For $\epsilon>0$, let N be sufficiently large so that $|x_n-x_{n-2}|<\epsilon$ for all $n\geqq N$. Note that for any $n>N$,

$$\begin{aligned} x_n - x_{n-1} = (x_n - x_{n-2}) - (x_{n-1} - x_{n-3}) + (x_{n-2} - x_{n-3}) - \cdots \\ \pm (x_{N+1} - x_{N-1}) \mp (x_N - x_{N-1}). \end{aligned}$$

Thus $|x_n-x_{n-1}| \leqq (n-N)\epsilon+|x_N-x_{N-1}|$ and $\lim_{n\to\infty} (x_n-x_{n-1})/n=0$.

A-5 (11). Since parallel cross sections of the ellipsoid are always similar ellipses, any circular cross section can be increased in size by taking a parallel cutting plane passing through the center. Every plane through (0, 0, 0) which makes a circular cross section must intersect the y-z plane. But this means that a diameter of the circular cross section must be a diameter of the ellipse $x=0$, $y^2/b^2+z^2/c^2=1$. Hence the radius of the circle is at most b. Similar reasoning with the x-y plane shows that the radius of the circle is at least b, so that any circular cross section formed by a plane through (0, 0, 0) must have radius b, and this will be the required maximum radius. To show that circular cross sections of radius b actually exist, consider all planes through the y-axis. It can be verified that the two planes given by $a^2(b^2-c^2)z^2=c^2(a^2-b^2)x^2$ give circular cross sections of radius b.

A-6 (48). Let x be selected from $[0, L_1]$, y from $[0, L_2]$, z from $[0, L_3]$, and assume $L_3\geqq L_2\geqq L_1$. Let $X=\min(x, y, z)$.

$$L_1L_2L_3E[X] = \int_0^{L_1}\int_0^{L_2}\int_0^{L_3} X\,dz\,dy\,dx$$

$$= \int_0^{L_1}\int_0^{L_2}\left\{\int_0^{\mu} z\,dz + \int_{\mu}^{L_3}\mu\,dz\right\}dy\,dx, \quad \text{where } \mu = \min(x, y),$$

$$= \int_0^{L_1}\int_0^{L_2}\left\{L_3\mu - \tfrac{1}{2}\mu^2\right\}dy\,dx$$

$$= \int_0^{L_1}\left\{\int_0^{x}(L_3y - \tfrac{1}{2}y^2)dy + \int_x^{L_2}(L_3x - \tfrac{1}{2}x^2)dy\right\}dx$$

$$= \cdots = \frac{1}{2}L_1^2L_2L_3 - \frac{1}{6}L_1^3(L_2 + L_3) + \frac{1}{12}L_1^4.$$

Alternate Solution: For $0 \leqq a \leqq L$,

$$\begin{aligned} P(X \leqq a) &= P(x \leqq a) + P(y \leqq a) + P(z \leqq a) - P(x \leqq a)P(y \leqq a) \\ &\quad - P(x \leqq a)P(z \leqq a) - P(y \leqq a)P(z \leqq a) \\ &\quad + P(x \leqq a)P(y \leqq a)P(z \leqq a) \\ &= \frac{a}{L_1} + \frac{a}{L_2} + \frac{a}{L_3} - \left(\frac{a^2}{L_1L_2} + \frac{a^2}{L_2L_3} + \frac{a^2}{L_3L_1}\right) + \frac{a^3}{L_1L_2L_3}\,. \end{aligned}$$

The answer follows easily from the formula

$$E[X] = \int_0^{L_1} a\,\frac{dP(X \leqq a)}{da}\,da.$$

Comment: The first solution was presented by Robert Oliver and the alternate solution was presented by Jockum Aniansson.

SOLUTIONS. PART B

B-1 (75). Let

$$a_n = \frac{1}{n^4}\prod_{i=1}^{2n}(n^2 + i^2)^{1/n}.$$

Then

$$\log a_n = \frac{1}{n}\sum_{i=1}^{2n}\log\left(1 + \frac{i^2}{n^2}\right),$$

and

$$\operatorname*{limit}_{n\to\infty}\log a_n = \int_0^2 \log(1 + x^2)dx = 2\log 5 - 4 + 2\arctan 2.$$

B-2 (598). Let $P(t) = at^3 + bt^2 + ct + d$. The equation

$$\frac{1}{2T}\int_{-T}^{T}P(t)dt = \frac{1}{2}\left\{P(t_1) + P(t_2)\right\}$$

is satisfied for all values of a, b, c, and d if and only if $t_2=-t_1=\pm T/\sqrt{3}$. If $T=3$ hrs, $T/\sqrt{3}\approx 1$ hr, 43.92 min. Therefore, in the case considered, the critical times are 1 hour 44 minutes each side of noon.

B-3 (167). Let $y_n \to y$ with $(x_n, y_n) \in S$ for all n. The Bolzano-Weierstrass Theorem implies that a subsequence $x_{k(n)} \to x$. Then $y_{k(n)} \to y$ and since S is closed, $(x, y) \in S$. Thus y is in the projection of S on the y-axis.

B-4 (58). Converting units to feet and seconds, we have $0 \leqq v(t) \leqq 132$ for all $t \in [0, 60]$. Suppose $|v'(t)| < 6.6$ for all $t \in [0, 60]$. Then $v(t) = \int_0^t v' < 6.6t$, and $v(t) = \int_t^{60} -v' < 6.6(60-t)$ for all $t \in [0, 60]$. Thus

$$5280 = \int_0^{60} v(t)dt < \int_0^{60} \min\{6.6t, 6.6(60-t), 132\}dt.$$

This last integral is the area under a trapezoid and equals the value 5280, which is a contradiction.

Comment: Several students made the tacit assumption that the optimum graph for $v(t)$ was the trapezoid referred to above, but failed to give explicit justification.

B-5 (102). Clearly u_n is continuous. So, if F is continuous, then $u_n \circ F$ is the composition of continuous functions and hence is continuous. Conversely, suppose $u_n \circ F$ is continuous for all n. To prove F is continuous it is enough to show $F^{-1}[(a, b)]$ is open for every bounded interval (a, b). Let $n > \max(|a|, |b|)$. Then $u_n^{-1}[(a, b)] = (a, b)$ so

$$F^{-1}[(a, b)] = F^{-1}[u_n^{-1}\{(a, b)\}] = (u_n \circ F)^{-1}[(a, b)],$$

which is an open set by the continuity of $u_n \circ F$.

B-6 (6). Since the quadrilateral is circumscribable, $a+c=b+d$. Let k be the length of a diagonal and angles α and β selected so that $k^2=a^2+b^2-2ab\cos\alpha = c^2+d^2-2cd\cos\beta$. If we subtract $(a-b)^2=(c-d)^2$, we obtain

$$2ab(1-\cos\alpha) = 2cd(1-\cos\beta). \tag{1}$$

From $A = \frac{1}{2}ab\sin\alpha + \frac{1}{2}cd\sin\beta = \sqrt{abcd}$,

$$4A^2 = 4abcd = a^2b^2(1-\cos^2\alpha) + c^2d^2(1-\cos^2\beta) + 2abcd\sin\alpha\sin\beta.$$

Using (1) twice on the right hand side,

$$4abcd = ab(1+\cos\alpha)cd(1-\cos\beta) + cd(1+\cos\beta)ab(1-\cos\alpha) + 2abcd\sin\alpha\sin\beta.$$

On simplifying, $4=2-2\cos(\alpha+\beta)$, which implies that $\alpha+\beta=\pi$ and so the quadrilateral is cyclic.

Comment: This elegant solution was presented by Robert Oliver. He was the only student to receive a perfect score on this problem.

Acknowledgments

The Director would like to acknowledge the assistance of the Questions Committee and the graders in preparing the above solutions and acknowledge the services of the following persons, who were graders for the competition: D. M. Bloom, L. R. Bragg, R. C. Busby, C. Chipman, C. V. Coffman, E. C. Curtis, R. A. DeVore, R. M. Dudley, W. Emerson, D. J. Eustice, J. Froemke, R. Gambill, M. Hausner, W. S. Loud, D. G. Malm, J. M. Martin, E. A. Nordhaus, R. Pollack, D. Rosen, P. Sally, I. Schochetman, M. Shanks, W. Snow, M. L. Tomber, J. P. Williams, E. T. Wong.

MATHEMATICAL NOTES

EDITED BY DAVID DRASIN

Manuscripts for this Department should be sent to Robert Gilmer, Department of Mathematics, Florida State University, Tallahassee, FL 32306.

A FUNCTIONAL ANALYTIC PROOF OF ROUCHÉ'S THEOREM

D. VAN DULST, University of Maryland

In this note we show how Rouché's theorem (cf. [3]) can be derived from an index theorem for linear operators in Banach spaces which is due to T. Kato [2]. Using this approach one needs only two facts concerning analytic functions, namely the possibility of power series expansion and the maximum modulus principle. No integration is involved in this proof.

1. Preliminaries. (We refer to [1] for more details.) Let $T:X\to Y$ be a linear operator from a Banach space X into a Banach space Y. We use the following standard notations:

$$\alpha(T) = \dim N(T), \qquad \text{where } N(T) \text{ is the nullspace of } T.$$
$$\beta(T) = \dim [Y/R(T)], \qquad \text{where } R(T) \text{ is the range of } T.$$

If at least one of the numbers $\alpha(T)$ and $\beta(T)$ is finite, then T has an **index** $\kappa(T)$ (possibly infinite), defined by $\kappa(T)=\alpha(T)-\beta(T)$.

The following result is due to T. Kato (cf. [2] for a more general version).

THEOREM (T. Kato). *Let T and B be bounded linear operators from a Banach space X into a Banach space Y. Suppose that T has an index and that $R(T)$ is closed. If there exists a finite constant a such that*

$$\|Bx\| \leqq a\|Tx\| \quad \textit{for all } x \in X, \tag{1}$$

then for all $\lambda\in\mathbf{C}$ such that

$$|\lambda| < 1/a \tag{2}$$

we have

AN INEQUALITY INVOLVING THE AREA OF TWO TRIANGLES

L. CARLITZ, Duke University

Let a, b, c denote the sides of the triangle ABC and let a', b', c' denote the sides of the triangle $A'B'C'$. Let F, F' denote the respective areas. Pedoe has proved that

$$a^2(-a'^2+b'^2+c'^2)+b^2(a'^2-b'^2+c'^2)+c^2(a'^2+b'^2-c'^2) \geqq 16FF', \tag{1}$$

with equality if and only if the triangles ABC, $A'B'C'$ are similar. For related results and references see [1].

It may be of interest to give a simple algebraic proof of (1). We have

$$16F^2 = 2b^2c^2+2c^2a^2+2a^2b^2-a^4-b^4-c^4,$$
$$16F'^2 = 2b'^2c'^2+2c'^2a'^2+2a'^2b'^2-a'^4-b'^4-c'^4.$$

Put

$$\begin{aligned} x, y, z &= a^2, b^2, c^2; \quad x', y', z' = a'^2, b'^2, c'^2,\\ Q(x) &= 2yz+2xz+2xy-x^2-y^2-z^2,\\ Q(x') &= 2y'z'+2z'x'+2x'y'-x'^2-y'^2-z'^2,\\ Q(x,x') &= x(-x'+y'+z')+y(x'-y'+z')+z(x'+y'-z')\\ &= x'(-x+y+z)+y'(x-y+z)+z'(x+y-z). \end{aligned}$$

Then (1) becomes

$$Q^2(x,x') \geqq Q(x)Q(x'). \tag{2}$$

It can be verified that

$$Q^2(x,x')-Q(x)Q(x') = -4(VW+WU+UV), \tag{3}$$

where $U=yz'-y'z$, $V=zx'-z'x$, $W=xy'-x'y$. Since $xU+yV+zW=0$, we get

$$\begin{aligned} &-4xz(VW+WU+UV) = -4xzUV+4x(U+V)(xU+yV)\\ &= 4x^2U^2+4x(x+y-z)UV+4xyV^2 = [2xU+(x+y-z)V]^2+Q(x)V^2. \end{aligned}$$

Hence

$$xz[Q^2(x,x')-Q(x)Q(x')] = [2xU-(x-y-z)V]^2+Q(x)V^2. \tag{4}$$

Since $xz>0$, $Q(x)\geqq 0$, it is clear that (4) implies (2). Moreover, we shall have equality in (2) if and only if $U=V=W=0$, that is, if and only if

$$\frac{x'}{x}=\frac{y'}{y}=\frac{z'}{z}.$$

This evidently completes the proof of Pedoe's theorem.

Reference

1. D. Pedoe, Thinking geometrically, this MONTHLY, 77 (1970) 711–721.

A GENERALIZATION OF A THEOREM OF BOREL CONCERNING THE DISTRIBUTION OF DIGITS IN DYADIC EXPANSIONS

JÁNOS GALAMBOS, University of Ibadan, Nigeria

1. Introduction. Many classroom examples for sequences of independent random variables are provided by algorithms to expand real numbers into infinite series of rationals. The best known expansion is the decimal, or the more general q-adic expansion,

$$x = \sum_{k=1}^{+\infty} e_k q^{-k}, \qquad 0 < x < 1, \tag{1}$$

where $q > 1$ is an integer and e_k can take the values $0, 1, \cdots, q-1$. Here the digits e_k are obtained by the following algorithm. First let $x = x_1$; then e_1 is the largest integer less than qx_1, so that

$$e_1 < qx_1 \leqq e_1 + 1.$$

We then write $x_2 = qx_1 - e_1$. Similarly, if x_k has been defined, let e_k be the largest integer less than qx_k, which can equivalently be defined as the unique integer satisfying

$$e_k < qx_k \leqq e_k + 1, \tag{2a}$$

and then let

$$x_{k+1} = qx_k - e_k, \tag{2b}$$

and continue the above procedure. Note that e_k slightly differs from the integer part $[qx_k]$ of qx_k, since if qx_k is an integer, then $e_k = qx_k - 1$, otherwise $e_k = [qx_k]$. This choice guarantees that infinitely many e_k differ from 0. The digits e_k are evidently functions $e_k(x)$ of x, and are random variables on the probability space $(\Omega, \mathfrak{A}, P)$, where $\Omega = (0, 1)$, P is Lebesgue measure, and $\mathfrak{A}$ is the set of Lebesgue measurable subsets of $(0, 1)$. (Throughout the present paper, this specific probability space is used.) It was observed by Borel [2] that the random variables $e_k(x)$ are independent and identically distributed, taking any of the values $0, 1, \cdots, q-1$ with the same probability $1/q$. This fact, through the strong law of large numbers [3, p. 244], implies that if $q = 2$, the relative frequency of the digits 0 and 1 tends to $\frac{1}{2}$ for almost all x.

Let $1 < q < 2$ in (1) and (2a, b). The random variables e_k, determined in (2a) and hence taking the values 0 or 1, are no longer independent, and it is a very deep theorem that the relative frequencies of 0 and 1 again have limits, the same for almost all x. This fact was proved by A. Rényi [7] by making use of deep analytic and probabilistic tools. Another possibility to get information about the ones and zeros among the digits e_k in (1) is an approach suggested by A. Oppenheim (personal communication). Dropping the terms in (1) where $e_k = 0$, we get

$$x = \sum_{k=1}^{+\infty} q^{-n_k}, \tag{3}$$

where $n_1 < n_2 < \cdots$ are integers and are obviously functions $n_k(x)$ of x. The event (set) $\{n_1=4\}$, say, means that $\{e_1=0,\ e_2=0,\ e_3=0,\ e_4=1\}$, and in general, the e_k and the n_k are related by

$$\{\text{at least } k \text{ out of } e_1, e_2, \cdots, e_N \text{ equal } 1\} = \{n_k \leqq N\}. \tag{4}$$

Because of the strong dependence of the e's, it is difficult to evaluate the probability of the left hand side of (4); it is, however, possible to evaluate the probability of the right hand side and, as a corollary, it is obtained in [5] that if there is a positive integer a with $q^{a+1}-q^a=1$, then the random variables n_1, n_2-n_1, $n_3-n_2, \cdots$ are independent. We can not, however, claim that we therefore have a simple approach to evaluate $P(n_k<N)$ if $q^{a+1}-q^a=1$, with $a \geqq 1$ an integer, since the theorem from which the corollary mentioned is deduced is fairly complicated. In this paper I shall present a very simple proof of this statement.

MAIN THEOREM. *Let $a \geqq 0$ be an integer, and let $1<q\leqq 2$ be the (unique) solution of $z^{a+1}-z^a=1$. Then the random variables $m_k=n_k-n_{k-1}$, $k=1, 2, \cdots$, (where $n_0=0$), are independent, and for $k \geqq 2$,*

$$P(m_k = t) = \begin{cases} q^{-t} & \text{for } t \geqq a+1 \\ 0 & \text{otherwise.} \end{cases}$$

As an application of the main theorem and of (4), we shall evaluate the limits of the relative frequency of ones and zeros among the e_k. The results are contained in Theorems 1–3 and the proof is split into lemmas. Out of the lemmas, Lemma 1 itself shows a very interesting characteristic of the expansions (1) and (3). (See the remark after its proof.)

There are two other examples from the field of representation of real numbers by infinite series of rationals that may be interesting for classroom presentation. One is found in [6], where independent random variables are obtained that are 'uniformly distributed on the infinite sequence of positive integers,' and the other one is in the last section of [4]; it provides a sequence of random variables with infinite expectation, the sum of which is asymptotically $n \log n$ in probability, but this can not be replaced by an asymptotic law with probability unity. For these, the original proofs themselves are simple.

For further references in the field of probabilistic aspects of rational approximations see [5].

2. The Distribution of n_k. There is a direct algorithm to obtain (3). Let $x=x_1$, let n_1 be the smallest positive integer for which q^{-n_1} is smaller than x_1, and let $x_2=x_1-q^{-n_1}$. If x_k has been defined, let n_k be the smallest integer for which $q^{-n_k}<x_k$, and let $x_{k+1}=x_k-q^{-n_k}$, i.e., in short, the algorithm for obtaining the

sequence n_k is

$$x = x_1, \qquad q^{-n_k} < x_k \leqq q^{-n_k+1}, \qquad x_{k+1} = x_k - q^{-n_k}. \tag{5}$$

In the sequel we assume that q is the (only) root in (1, 2) of $z^{a+1}-z^a=1$, where $a \geqq 0$ is an integer.

LEMMA 1. *If* $m_k = n_k - n_{k-1}$ *for* $k=2, 3, \cdots$, *then* $m_k \geqq a+1$.

Proof. By the repeated application of (5), we have

$$q^{-n_k} + q^{-n_{k+1}} < x \leqq q^{-n_k+1}$$

for any $k>0$, hence

$$q^{-n_k} + q^{-n_k-m_{k+1}} < q^{-n_k+1},$$

i.e., $1+q^{-m_{k+1}}<q$. Since $q=1+q^{-a}$, the lemma is established.

It is worthwhile to look further at the statement of Lemma 1. It implies that if $a \geqq 1$, then $m_k \geqq 2$, i.e., a digit one in (1) is necessarily followed by a digit zero, which in fact already explains that the e_k can not be independent of each other. It also tells us that if $q \in (1, 2)$ is the root of $z^2-z=1$, then although

$$1 = q^{-1} + \sum_{k=4}^{+\infty} q^{-k},$$

it is not the expansion (1) and (2) of unity, since here $e_k = e_{k+1} = 1$ for all $k \geqq 4$, hence contradicting our result in Lemma 1. This shows that in order to speak of the expansion of real numbers in a given form, the algorithm used to obtain that form must be specified.

LEMMA 2.

$$P(n_1 = j_1, n_2 = j_2, \cdots, n_k = t) = \begin{cases} q^{-t+1}(1 - 1/q) & \text{if } j_s - j_{s-1} > a \quad \text{for } s \geqq 2 \\ 0 & \text{otherwise.} \end{cases}$$

Proof. In view of (5),

$$\{n_1 = j_1, n_2 = j_2, \cdots, n_k = t\} = \bigcap_{s=1}^{k} (A_s, B_s),$$

where $A_s = q^{-j_1} + q^{-j_2} + \cdots + q^{-j_s}$ and $B_s = A_{s-1} + q^{-j_s+1}$, with $j_k = t$. On the other hand we shall see that

$$\bigcap_{s=1}^{k} (A_s, B_s) = (A_k, B_k), \tag{6}$$

and therefore $P(n_1=j_1, \cdots, n_k=t) = B_k - A_k = q^{-t+1}(1-1/q)$, which will yield the lemma (the second part of the last statement is an immediate consequence of Lemma 1.) In order to prove (6), note first that $A_s < A_{s+1}$ for all s. On the other hand, $B_{s+1} \leqq B_s$, since

$$B_s - B_{s+1} = q^{-j_s+1} - q^{-j_s} - q^{-j_{s+1}+1} = q^{-j_s+1}(1 - q^{-1} - q^{-(j_{s+1}-j_s)}),$$

and the multiplier of q^{-j_s+1} is nonnegative in view of the assumptions $1-q^{-1}-q^{-a-1}=0$ and $j_{s+1}-j_s \geqq a+1$. Thus (6) and hence the lemma is proved.

Note that for arbitrary q (6) is not valid, which makes the investigation in the general case much more complicated.

LEMMA 3.

$$P(n_{k+1} = t + 1) = \binom{t - ka}{k} \frac{1}{q^t}\left(1 - \frac{1}{q}\right).$$

Proof. The law of total probability implies that

$$P(n_{k+1} = t + 1) = \sum P(n_1 = j_1, \cdots, n_k = j_k, n_{k+1} = t + 1), \tag{7}$$

where the summation is for all k-vectors $(j_1, \cdots, j_k)$. By Lemma 2, the nonzero terms in (7) are all equal to $q^{-1}(1-1/q)$, independent of the values of $(j_1, \cdots, j_k)$, hence we have to count the number of nonzero terms in (7). By another appeal to Lemma 2, the number of nonzero terms in (7) is the number of $(k+1)$-vectors with last component $t+1$, and for which the differences between the coordinates are at least $a+1$. This was found in [1] and [5] to be the binomial coefficient with parameters $t-ka$ and k, and thus the proof of Lemma 3 is completed.

LEMMA 4. *The conditional probability satisfies*

$$P(n_{k+1} = t + j \mid n_k = j) = q^{-t} \qquad \textit{for } t \geqq a + 1.$$

Proof. Using the same argument as in the proof of Lemma 3, we get

$$P(n_{k+1} = t + j,\, n_k = j) = \binom{j - (k-1)a - 1}{k - 1} \frac{1}{q^{t+j-1}}\left(1 - \frac{1}{q}\right) \qquad \text{if } t \geqq a + 1. \tag{8}$$

By the definition of conditional probabilities,

$$P(n_{k+1} = t + j \mid n_k = j) = P(n_{k+1} = t + j,\, n_k = j)/P(n_k = j),$$

and thus Lemma 3 and (8) yield Lemma 4.

THEOREM 1. *For* $k \geqq 1$,

$$P(n_{k+1} - n_k = t) = \begin{cases} q^{-t} & \textit{for } t \geqq a + 1 \\ 0 & \textit{otherwise.} \end{cases}$$

Proof. By the total probability rule [3, p. 106],

$$P(n_{k+1} - n_k = t) = \sum_{j=ka}^{+\infty} P(n_{k+1} = t + j \mid n_k = j) P(n_k = j),$$

which by Lemmas 3 and 4 gives

$$P(n_{k+1} - n_k = t) = \frac{1}{q^t} \sum_{j=ka}^{+\infty} P(n_k = j) = \frac{1}{q^t},$$

as stated. The second part is immediate from Lemma 1.

THEOREM 2. *The random variables $m_1 = n_1$, $m_k = n_k - n_{k-1}$ for $k \geqq 2$ are independent.*

Proof. Evidently

$$P(m_1 = t_1, \cdots, m_k = t_k) = P(n_1 = t_1, n_2 = t_1 + t_2, \cdots, n_k = t_1 + \cdots + t_k),$$

which by Lemma 2 equals 0 if there is an s such that $t_s \leqq a$, and equals

$$q^{-t_1 - \cdots - t_k - 1}(1 - 1/q)$$

otherwise. Hence Lemma 2 with $k = 1$ and Theorem 1 terminate the proof.

Theorems 1 and 2 constitute the Main Theorem stated in the introduction. As an application, let us prove the following theorem.

THEOREM 3. *Let $F(N, 1, x)$ and $F(N, 0, x)$ denote the number of $k \leqq N$ such that $e_k(x) = 1$ and $e_k(x) = 0$, respectively, in* (1). *Then for almost all x,*

$$\lim_{N=+\infty} \frac{F(N, 1, x)}{N} = \frac{q - 1}{aq - a + q}, \qquad \lim_{N=+\infty} \frac{F(N, 0, x)}{N} = \frac{aq - a + 1}{aq - a + q}.$$

Proof. By Theorems 1 and 2 and by the strong law of large numbers [**3**, p. 244], for almost all x,

$$\frac{n_k}{k} = \frac{(m_1 + \cdots + m_k)}{k} \to \sum_{t=a+1}^{+\infty} tq^{-t}$$

$$= \frac{1}{q^{a+1}} \sum_{s=0}^{+\infty} \frac{s+1}{q^s} + a \sum_{t=a+1}^{+\infty} \frac{1}{q^t} = \frac{1}{q^{a+1}} \frac{1}{(1 - 1/q)} + \frac{a}{q^{a+1}} \frac{1}{(1 - 1/q)}$$

$$= \frac{aq - a + q}{q - 1}.$$

But $\{F(N, 1, x) \leqq k\} = \{n_k \geqq N\}$, hence the first limit in Theorem 3 is proved; since the two limits add to one, the proof is completed.

Note that for $a = 0$, $q = 2$, the limits equal $\frac{1}{2}$, thus giving back Borel's result.

The author is now at Temple University, Philadelphia.

References

1. H. D. Abramson, On selecting separated objects from a row, this MONTHLY, 76 (1969) 1130–31.

2. E. Borel, Sur les probabilités dénombrables et leurs applications arithmétiques, Rend. Circ. Mat. Palermo, 26 (1909) 247–271.

3. W. Feller, An Introduction to Probability Theory and its Applications, 2nd ed., vol. I, Wiley, New York, 1957.

4. J. Galambos, The ergodic properties of the denominators in the Oppenheim expansion of real numbers into infinite series of rationals, Quart. J. Math., Oxford Ser. (2), 21 (1970) 177–191.

5. ———, Probabilistic theorems concerning expansions of real numbers, To appear in an Alfréd Rényi Memorial Volume to be published in Hungary.

6. A. Rényi, The distribution of digits in the Cantor series of real numbers (in Hungarian), Mat. Lapok, 7 (1956) 77–100.

7. ———, Representations for real numbers and their ergodic properties, Acta Math. Acad. Sci. Hungar., 8 (1957) 477–493.

RESEARCH PROBLEMS

EDITED BY RICHARD GUY

In this Department the Monthly presents easily stated research problems dealing with notions ordinarily encountered in undergraduate mathematics. Each problem should be accompanied by relevant references (if any are known to the author) and by a brief description of known partial results. Material should be sent to Richard Guy, Department of Mathematics, Statistics, and Computing Science, The University of Calgary, Calgary 44, Alberta, Canada.

A LOWER BOUND OF FIELDS DUE TO UNIT POINT MASSES

C. K. CHUI, Texas A&M University

Let U denote the open unit disc and T the unit circle. The complex conjugate of

$$S_n(z) = \sum_{k=1}^{n} \frac{1}{z - z_{n,k}},$$

where $z_{n,k} \in T$, $k=1, \cdots, n$, represents the gravitational (or electrostatic) field at the point z due to unit point masses (or charges) at the points $z_{n,k}$ on T. [Cf. 2.]

Suppose $z_{n,k} = e^{2\pi ik/n}$ and write $w_n = e^{2\pi i/n}$. Then

$$\sum_{k=1}^{n} \frac{1}{z - w_n^k} = \frac{-nz^{n-1}}{1 - z^n} \to 0$$

uniformly on each compact subset of U as $n \to \infty$. Hence, although the total mass on T is n, which tends to infinity, the field at each point $z \in U$ tends to zero. However, the average field strength in U due to unit masses at w_n^k is

$$\frac{1}{\pi} \int\int_U \left| \sum_{k=1}^{n} \frac{1}{(z - w_n^k)} \right| dxdy = \frac{1}{\pi} \int_0^{2\pi} \left\{ \int_0^1 \frac{nr^{n-1}}{\sqrt{1 - 2r^n \cos n\theta + r^{2n}}} rdr \right\} d\theta$$
$$= \frac{1}{\pi} \int_0^{2\pi} \left\{ \int_0^1 \frac{r^{1/n} dr}{\sqrt{1 - 2r \cos \theta + r^2}} \right\} d\theta,$$

which converges to the positive number

$$\frac{1}{\pi}\int_0^{2\pi}\int_0^1 \frac{dr d\theta}{\sqrt{1-2r\cos\theta+r^2}}$$

as $n\to\infty$. Hence there is a number $m>0$ such that

$$\frac{1}{\pi}\iint_U \left|\sum_{k=1}^n 1/(z-\overset{k}{w_n})\right| dxdy \geqq m$$

for all $n=1, 2, \cdots$.

CONJECTURE. For all $z_{n,k}\in T$, $k=1, \cdots, n$, $n=1, 2, \cdots$,

$$\iint_U \left|\sum_{k=1}^n 1/(z-z_{n,k})\right| dxdy \geqq \iint_U \left|\sum_{k=1}^n 1/(z-\overset{k}{w_n})\right| dxdy.$$

That is, we conjecture that the average field strength in U due to unit point masses on T is the smallest if these point masses are equally spaced on T. Hence, this average field strength in U is uniformly bounded below by $m>0$, no matter where the n point masses are placed on T.

As an application, we consider the following approximation problem. If f is holomorphic in U, then it is not difficult to prove [3] that there exist functions $S_n(z)=\sum_{k=1}^n 1/(z-z_{n,k})$, $z_{n,k}\in T$, $k=1, \cdots, n$ and $n=1, 2, \cdots$, such that $S_n\to f$ uniformly on each compact subset of U. Let H be the Banach space of all holomorphic functions f in U with finite area norm:

$$\|f\| = \frac{1}{\pi}\iint_U |f(x+iy)|\, dxdy < \infty.$$

By successive applications of the Hahn-Banach, Riesz Representation, and Stokes Theorems [1], it is quite easy to show that the collection of all functions

$$t_n(z) = \sum_{k=1}^n a_{n,k}/(z-z_{n,k}),$$

where the $a_{n,k}$ are real numbers and $z_{n,k}\in T$, is dense in H. However, if the above conjecture is true, then the class of all functions $S_n(z)=\sum_{k=1}^n 1/(z-z_{n,k})$, $z_{n,k}\in T$, cannot be dense in H. In particular, there would exist a function f in H which cannot be approximated in H by functions S_n, although f can always be approximated by S_n uniformly on each compact subset of U.

References

1. L. Bers, An approximation theorem, J. Analyse Math., 14 (1965) 1–4.
2. J. Korevaar, Limits of polynomials whose zeros lie in a given set, Proc. Symposia Pure Math. Vol. XI, Amer. Math. Soc., (1968) 261–267.
3. G. R. MacLane, Polynomials with zeros on a rectifiable Jordan curve, Duke Math. J., 16 (1949) 461–477.

CLASSROOM NOTES

EDITED BY DAVID DRASIN

Manuscripts for this Department should be sent to Robert Gilmer, Department of Mathematics, Florida State University, Tallahassee, FL 32306.

ON YOUNG'S INEQUALITY

F. CUNNINGHAM, JR., Bryn Mawr College and
NATHANIEL GROSSMAN, University of California, Los Angeles

The inequality of the title is

$$ab \leqq \int_0^a f(x)dx + \int_0^b f^{-1}(y)dy \tag{1}$$

for $a \geqq 0$, $b \geqq 0$, where f is strictly increasing and continuous (so that f^{-1} exists); equality holding in (1) if and only if $b = f(a)$. (We assume throughout for convenience that both f and f^{-1} have $[0, +\infty)$ for domain, modifications to fit other cases being obvious.) There seems to be a prevalent belief (cf. [1]) that while the truth of this theorem is easily grasped from a figure, an analytic proof is necessarily difficult. The message of this note is that a proof based on a simple manipulation of Riemann sums for both integrals is short, builds squarely on the geometric hint, and permits a natural and seemingly unimprovable generalization. It seems certain that both proof and generalization are widely known. This note is submitted not with any claim of originality, but to correct what appears to be a gap in the literature. We first give the proof of the theorem as already stated, and generalize it afterwards.

Assume first that $b = f(a)$; we shall prove equality in (1) in this case. Since monotone functions are Riemann integrable, given $\epsilon > 0$ we can find nondecreasing step functions f_1 and f_2 such that $f_1(x) \leqq f(x) \leqq f_2(x)$ on $[0, a]$ and

$$\int_0^a f_2(x)dx - \int_0^a f_1(x)dx < \epsilon.$$

Now construct step-functions g_1 and g_2 by "inverting" f_1 and f_2 as follows. (g_1 and g_2 will be pseudo-inverses of f_1 and f_2 in the sense defined below.) The interval $[0, a]$ is partitioned into intervals $[x_{i-1}, x_i]$, $i = 1, \cdots, n$, $x_0 = 0$, $x_n = a$, on the interior of each of which f_1 has a constant value y_i. Set $y_0 = 0$ and $y_{n+1} = b$. Then $[0, b]$ is partitioned into intervals $[y_{i-1}, y_i]$, $i = 1, \cdots, n+1$. Let g_1 have the constant value x_{i-1} on $[y_{i-1}, y_i]$, and set $g_1(b) = a$. Similarly, make g_2 from f_2. Now it is easily verified that $g_2(y) \leqq f^{-1}(y) \leqq g_1(y)$ on $[0, b]$. Moreover, a calculation with finite sums yields immediately

$$\int_0^a f_1(x)dx + \int_0^b g_1(y)dy = ab = \int_0^a f_2(x)dx + \int_0^b g_2(y)dy.$$

From these one easily deduces

$$\left| ab - \int_0^a f(x)dx - \int_0^b f^{-1}(y)dy \right| < 2\epsilon,$$

whence the alleged equality since ϵ is arbitrary.

Next suppose $b > f(a) = b'$. Then $f^{-1}(y) > a$ for all y in $(b', b]$, whence

$$\int_{b'}^b f^{-1}(y)dy > a(b - b').$$

Applying the first part of the theorem with b replaced by b', we obtain

$$\begin{aligned}\int_0^b f^{-1}(y)dy &= \int_0^{b'} f^{-1}(y)dy + \int_{b'}^b f^{-1}(y)dy \\ &> ab' - \int_0^a f(x)dx + a(b - b') \\ &= ab - \int_0^a f(x)dx,\end{aligned}$$

so that strict inequality holds in (1). The remaining case $b < f(a)$ is treated similarly, interchanging the roles of f and f^{-1}, to complete the proof.

Generalizing, we drop the requirements that f be *strictly* increasing and continuous, retaining only the assumption that f is nondecreasing on $[0, +\infty)$. To serve in place of f^{-1}, which now need not make sense, we use a pseudo-inverse defined as follows. For each $y \geqq 0$, define

$$x_L(y) = \sup\{x \geqq 0 \mid f(x) < y\} \quad \text{and} \quad x_R(y) = \inf\{x \geqq 0 \mid f(x) > y\},$$

except that we set $x_L(0) = 0$. Then $x_L(y) \leqq x_R(y)$, and any function g on $[0, +\infty)$ such that $x_L(y) \leqq g(y) \leqq x_R(y)$ for all $y \geqq 0$ is called a *pseudo-inverse* of f. Any such g is nondecreasing, and f is a pseudo-inverse of g. If f is strictly increasing, it has only one pseudo-inverse, which is f^{-1} if f is also continuous. When more than one pseudo-inverse exist, any two of them agree except on a countable set, so their integrals are the same.

THEOREM. *Let f and g be nondecreasing functions on $[0, +\infty)$ such that g is a pseudo-inverse of f, and let $a \geqq 0$, $b \geqq 0$. Then*

$$ab \leqq \int_0^a f(x)dx + \int_0^b g(y)dy, \tag{1'}$$

equality holding if and only if

$$f(a-) \leqq b \leqq f(a+). \tag{2}$$

The proof is the same with obvious small changes. Read g for f^{-1} throughout. The first case is when (2) holds. Then redefine f at a to make $f(a) = b$ and proceed as before. The second and third cases are $b > f(a+) = b'$ and $b < f(a-)$ respectively.

REMARK. It is possible for equality to hold in (1′) without either $b=f(a)$ or $a=f(b)$.

Reference

1. J. B. Diaz and F. T. Metcalf, An analytic proof of Young's inequality, this MONTHLY. 77 (1970) 603–609.

AN ELEMENTARY PROOF OF THE LEBESGUE DECOMPOSITION THEOREM

J. YAM TING WOO, University of California, Berkeley

In this note, we shall give a proof of the Lebesgue Decomposition Theorem that uses nothing from measure theory beyond the definitions needed to state the theorem. In particular, we do not need the Radon-Nikodym Theorem, and the Axiom of Choice is not explicitly used. We follow the notations of [1].

THEOREM. *Let μ be an arbitrary positive measure and λ a bounded positive measure on a σ-algebra $\mathfrak{M}$ in a set X. There is a unique pair of positive measures λ_1 and λ_2 on $\mathfrak{M}$ such that $\lambda=\lambda_1+\lambda_2$, $\lambda_1\perp\mu$, $\lambda_2\ll\mu$.*

Proof. Uniqueness is trivial. Existence is proved by contradiction. Assume that whenever we have λ_1, λ_2 such that $\lambda=\lambda_1+\lambda_2$ and $\lambda_1\perp\mu$, then there exists $E\in\mathfrak{M}$ such that $\lambda_2(E)>0$ and $\mu(E)=0$.

Then for all $A\in\mathfrak{M}$ such that $\mu(X-A)=0$, there exists $B\subset A$, $B\in\mathfrak{M}$ such that $\mu(B)=0$ and $\lambda(B)>0$. For consider the positive measures λ_A and λ_{X-A} defined by

$$\lambda_A(E) = \lambda(E\cap A),\ \lambda_{X-A}(E) = \lambda(E\cap(X - A))$$

for all $E\in\mathfrak{M}$. Then $\lambda=\lambda_A+\lambda_{X-A}$ and $\lambda_{X-A}\perp\mu$. So there exists an $E\in\mathfrak{M}$ such that $\lambda_A(E)>0$ and $\mu(E)=0$. We take $A\cap E$ as B and the result follows.

Now let $\mathcal{C}=\{E\in\mathfrak{M}\mid\lambda(E)>0,\ \mu(E)=0\}$. From the previous paragraph, with $A=X$, we see that $\mathcal{C}\neq\emptyset$. Let $T=\{\lambda(E)\mid E\in\mathcal{C}\}$. T is nonempty and bounded above by $\lambda(X)$. Let $\{\lambda(E_n)\}$ be a sequence in T converging to sup T, and let $E=\bigcup E_n$. Then $\lambda(E)\geqq\lambda(E_n)$ for all n, and so $\lambda(E)\geqq\sup T>0$. On the other hand, $\mu(E)\leqq\sum\mu(E_n)=0$. Thus $E\in\mathcal{C}$ and $\lambda(E)=\sup T$.

Now $\mu(E)=0$ implies $\mu(X-(X-E))=0$, so there exists an $F\in\mathfrak{M}$ such that $F\subset X-E$, $\lambda(F)>0$, and $\mu(F)=0$. Consider $E\cup F$. The union is disjoint, so $\lambda(E\cup F)=\lambda(E)+\lambda(F)>\lambda(E)=\sup T>0$, and $\mu(E\cup F)=\mu(E)+\mu(F)=0$. This implies $E\cup F\in\mathcal{C}$, contradicting the definition of sup T.

REMARK. We do not place any condition on μ. We can also extend the result to the case where λ is σ-finite, as outlined in [1].

Reference

1. Walter Rudin, Real and Complex Analysis, McGraw-Hill, New York, 1966.

MATHEMATICAL EDUCATION

EDITED BY J. G. HARVEY AND M. W. POWNALL

Material for this Department should be sent to either of the editors: J. G. Harvey, Department of Mathematics, University of Wisconsin, Madison, WI 53706; M. W. Pownall, Department of Mathematics, Colgate University, Hamilton, NY 13346.

MATHEMATICS FOR THE UNDERGRADUATE IN THE SOCIAL SCIENCES

M. F. DACEY, Northwestern University

The social sciences are usually divided into the fields of anthropology, economics, geography, political science, psychology, and sociology. Demography is also included but is typically treated as a subdivision of sociology; linguistics may be included but history seldom is. Because of this diversity, generalizations about the social sciences are usually false; hopefully, though, the following generalizations are true in a sufficiently large number of cases that they convey some understanding of the position of mathematics in many undergraduate social science programs. One limitation to generalization is that a small number of social science courses encompass a multitude of mathematical topics at widely varying levels of sophistication. As a consequence, a list of specific mathematical topics is pertinent only to the contents of the social science courses on a few campuses. Moreover, a summary of the mathematical contents of social science courses fails to identify the role of mathematical training in the undergraduate education of social science students. In lieu of a list of specific mathematical topics, it seems more productive to examine some of the considerations that affect the participation of social science students in mathematics courses. Identification of some of the key variables may assist in the design of courses that are both useful and attractive to students and faculty of social science departments.

Characteristics of the Social Sciences.

1. The number of undergraduate social science courses using college level mathematics is small.
2. Most social scientists know and use very little mathematics that is above the level of first year calculus or introductory, pre-calculus statistics.
3. There is a small and slowly growing number of social scientists who have reasonably strong mathematical training.
4. Some subdivisions of the social sciences make extensive use of mathematics and statistics.

These differing levels of mathematics usage reflect two conditions. One is that at the research level the social sciences differ greatly in the use that is made of mathematics, and these differences are reflected in the composition of undergraduate courses. The other is the wide diversity of educational objectives of students in social science courses.

Various social science fields differ greatly in the use of mathematics and within each social science field there is also substantial variation. Mathematical economics and econometrics are largely the analysis of mathematical models,

but branches of economics such as public finance or labor economics make little or no use of mathematics. Similarly, learning theory includes many mathematical models but the average clinical psychologist knows and uses little mathematics. There is also variation at the departmental level. Many departments have no mathematically oriented social scientists, and few are exclusively mathematical. Consequently, nearly all social science departments have relatively many undergraduate level courses that are essentially nonmathematical and nonstatistical in content.

This diversity becomes significant when related to the interests of social science majors. In general, only a small proportion of students select a social science major because it satisfies career objectives or is an entree to a graduate program in the social sciences. A large group uses their undergraduate social science training as a basis for admission to professional schools such as law, business, or foreign service. Another group will enter occupations for which college is deemed desirable, though there is not explicit use of their undergraduate major. There is also a miscellaneous group without clearly defined professional, academic, or employment objectives that evidently majors in a social science because it is mildly interesting and not too difficult to get passing grades.

Because social science programs educate many students who do not have a strong commitment to the field, this service function presents the dilemma of how much mathematics should a social science department require of, say, a pre-law student. A complicating factor is that many of these students dislike mathematics and will select a sequence of courses that makes minimal use of mathematics. To the degree that course content reflects the need to maintain class enrollment, there is pressure to keep the mathematical prerequisites at a low level. A department that stresses the use of mathematics in its undergraduate courses risks decreasing enrollment of students whose educational objectives are served equally well by any one of several social sciences.

One consequence is that very little mathematics is used in most undergraduate courses, and the conditions of the academic marketplace are such that imminent change of this condition is highly unlikely. For the mathematics department that anticipates teaching mathematics to social science majors, the preceding conditions imply that the demand for such courses may not be large.

Mathematics Taught in Social Science Departments. The preceding observations do not imply that social scientists receive no mathematical training. Few escape mathematics completely. Many social science departments, possibly a majority, have a statistics sequence that is frequently required for majors. Typically, the statistics is pre-calculus and stresses computations and the use of statistical tables. In addition, a few social science courses utilize a wide variety of mathematical tools, including elementary calculus and differential equations, linear and matrix algebra, set theory, graph theory, linear programming, and simulation methods. The mathematics required for each course is frequently summarized in appendices of textbooks and, with the possible exception of basic calculus, the mathematics is usually taught in the course as need arises. Two

pervasive characteristics of these courses are the emphasis on computation and the rarity of mathematics prerequisites.

The mathematical content of these courses presents two challenges. Is it possible to construct a sequence of mathematics courses that encompasses these mathematical topics? If such a course sequence were developed and offered, would it be used by social science majors?

In the design of suitable mathematics courses, there is an immediate conflict between the theoretical orientation of mathematics and the emphasis in social science courses on computations and applications. A mathematics course that will attract social science students will be a compromise between these conflicting interests. The possibility of a compromise largely depends upon the interests and personalities of the particular individuals in the appropriate departments on a campus.

Another problem is that the theory underlying the diverse types of mathematics used in social science courses is too extensive to be accommodated within a single course or sequence. This means that a mathematics course is limited to a selection of topics. However, at many institutions each social science department will argue that their students require a particular mix of mathematical topics that is incompatible with the mixes required by other social sciences. Moreover, students will be motivated only by a course that emphasizes the particular applications in their major field.

Even if mathematics and social science departments are able to resolve these conflicting interests, specially designed courses may still fail to attract sizeable numbers of social science students. One reason is that mathematically oriented social scientists like to teach mathematically oriented social science courses and may resist innovations and courses that diminish their participation in the mathematical training of students. A related reason is that many of these social scientists feel they are more capable of teaching mathematics to social science students than are mathematicians. Also, many social science faculty feel that mathematicians are unresponsive to the needs of social science students, and thus are reluctant to encourage their students to take mathematics courses offered in mathematics departments.

Undergraduate Mathematics in Graduate Social Science Programs. If the preceding comments suggest that mathematics training is incompatible with the social sciences, this is not the case, because the reconciliation between mathematics and social sciences largely occurs in graduate training and research. Because this training usually involves undergraduate and beginning graduate level mathematics, it is pertinent to comment on the undergraduate training of these students and their training in mathematics as graduate students.

One small group is undergraduate social science majors who receive extensive mathematical training. A second group has undergraduate degrees in mathematics, physics, or engineering. Many mathematically oriented social science departments actively seek such students with the encouraging assurance that "undergraduate training in the social sciences is not required for admission."

The third group has had little college level mathematics, but is required to participate in a crash program that attempts to establish a basic competence in calculus and finite mathematics. Summer programs, programmed learning, and first semester math review courses are among the devices used. While mathematics departments may cooperate in these activities, most remedial mathematics is concentrated within the concerned social science department.

These three types of students predominate in programs that make intensive use of mathematics and encourage or require the continued acquisition of mathematical skills by students throughout their graduate program. Some of this continuing training is acquired in mathematics courses; however, most of it is obtained in social science courses or in applied mathematics courses in statistics, operations research, and similar departments. As a consequence, many graduate students acquire a reasonably high level of mathematical competence without exposure to the advanced undergraduate and beginning graduate level courses offered in the mathematics department.

Summary. The preceding comments illustrate the typical role of mathematics in the undergraduate education of social scientists. While the generalizations may not pertain to the situation on every campus, they suggest some of the variables that determine the types of mathematics programs that can effectively complement undergraduate social science programs.

One key variable is that the role of mathematics in a social science curriculum depends greatly on the social science faculties at each institution. The mathematics department that initiates the development of mathematics courses for social science students needs to be aware of the teaching and research interests that prevail at its institution. This suggests that there is no pressing need for a master plan labelled "Mathematics for Social Scientists." The CUPM report [1] has many fine attributes and identifies course contents that are suited to numerous social science programs; yet the report has been largely ignored by social scientists. Instead of being guided by a list of courses and mathematical topics, a more effective approach is for the mathematics department to explore the local situation and enlist the cooperation of relevant social science faculties in the joint development of courses appropriate to local needs and interests. Instead of assuming that there is one best sequence of courses, there is need to experiment with a variety of solutions, each adapted to the local situation and each attuned to the interests of the particular individuals in the concerned departments. Some of the preceding comments on the diversity of mathematical tools used in the social sciences and on the competitive position of social science courses in the marketplace of undergraduate education may help to identify the obstacles that confront the cooperative development of mathematical social science programs.

Because the education of many social scientists in undergraduate mathematics is acquired as a graduate student, at many universities, it may be fruitful to explore the possibility of offering remedial and accelerated undergraduate math courses designed for the special needs of research in the social sciences. A

recent report [2] suggests this is the level having the greatest current need for improvement in teaching methods. Chapter 9 of this study identifies some of the problems that inhibit implementation of such courses and suggests solutions that primarily involve the use of new methods in mathematical sciences teaching. However, the prime considerations are the particular circumstances at each institution and the establishment of rapport and cooperation between individuals in the concerned departments.

This is a revision of a paper given at the Section A session on Mathematics for the Undergraduate at the AAAS meetings, Chicago, 1970.

References

1. CUPM, Tentative Recommendations for the Undergraduate Mathematics Program of Students in the Biological, Management and Social Sciences, Berkeley, California, January, 1964.
2. William H. Kruskal, ed. Mathematical Sciences and Social Sciences, Prentice-Hall, Englewood Cliffs, 1970.

MATHEMATICS FOR THE UNDERGRADUATE PHYSICS STUDENT

MARY L. BOAS, DePaul University, Chicago

It has been traditional to have a course for graduate students in mathematical methods of physics. A similar course at the undergraduate (say sophomore-junior) level has never become very customary and I think this is too bad. At the present time the usual physics major "picks up" much of his knowledge of mathematical techniques in his physics courses. The few mathematics courses he does take go into much greater detail than he finds either interesting or useful. Meanwhile the topics not studied in mathematics classes must be mastered from sketchy introductions in the physics textbooks. If the student is not overwhelmed by the new physical ideas he is trying to master, he is likely to be by the combination of new physics and a new mathematical technique presented simultaneously.

The physics teacher's counter argument here is that the motivation of the immediate use of a technique is very important and that for this reason the mathematics should not be taught separately. Years of experience have convinced me otherwise. It is quite possible in a mathematical physics course to tell the students enough about some of the simpler applications of the topic at hand to keep up interest.

Mathematics teachers may argue that the science major should study mathematics with as much detailed proof as mathematics majors. My only objection to this is lack of time to cover so many courses. What alternative is there?

After about ten years' experience with a course at DePaul University and writing a book (*Mathematical Methods in the Physical Sciences*, Wiley, New York, 1966), which we now use as our text, I think I am in a position to say that a mathematical physics course of about one year's length covering the needed topics can be given. How? First, such a course must be taught as a scientist's methods course and not as a mathematicians' theory course. This does not have

to mean careless or incorrect teaching. I make a strong point of never teaching anything a student will ever have to unlearn. However, generality can be sacrificed with little loss at this stage. The special cases of theorems that are going to be needed are quite sufficient. Second, long and detailed proofs can be omitted; careful statements of the theorems can be substituted. With these two methods of condensation (limited generality and few proofs), we cover in one year the following topics: probability, infinite series, complex numbers, determinants and matrices, partial differentiation, multiple integrals, vector algebra and calculus, Fourier series, calculus of variations, transformation theory, diagonalization of matrices and applications, tensors, complex variables, special functions, Laplace and Fourier transforms, and partial differential equations.

This is a summary of a talk given at the AAAS meeting on December 28, 1970, as part of a symposium on mathematics in the undergraduate science program, jointly sponsored by CUPM. The full text appeared in the *Two Year College Mathematics Journal.*

THE SMALL GROUP-DISCOVERY METHOD AS APPLIED IN CALCULUS INSTRUCTION

NEIL DAVIDSON, University of Maryland

Is there a way to learn mathematics that involves student pacing, active learning, thinking, and interpersonal communication? These criteria can be met by combining a small group method [9] with discovery learning [1, 7, 8]. In the small group-discovery method, the student discusses challenging problems with a few of his colleagues. The author first used this method in 1967–1968 in a one-year pilot study with a freshman calculus class at the University of Wisconsin.

During the pilot study, the students learned mathematics by doing mathematics. They formulated some definitions, stated most of the theorems, proved the theorems, constructed some examples and counterexamples, and developed techniques for solving various classes of problems. The students sometimes learned new concepts by discussing open-ended questions. For example: How can you find the area under a given curve? What is meant by a tangent to a curve? What happens at a high or low point on a graph? What can you conclude if a function vanishes at the endpoints of an interval? Discussion of the questions led to the statement of definitions or theorems.

The students worked together at the blackboard in small groups, with three or four members per group. The teacher stated the following guidelines for the small groups: (1) The students work together cooperatively and achieve a group solution to the problem. (2) Everybody understands the solution before the group tackles a new problem. (3) People listen carefully and try, whenever possible, to build upon the ideas of others. (4) There is no specified leader of the group. (5) Everybody participates and no one dominates the discussion. (6) People take turns writing solutions on the board.

The teacher selected the content and arranged it for small group learning. Since existing textbooks were not suitable for that purpose, he prepared a set of dittoed notes. He sometimes talked with the entire class at the beginning of the period, usually for no more than five or ten minutes per day. During these brief class discussions he presented new concepts, raised questions for investigation, proposed problems, and so forth.

The teacher spent most of the class period with the small groups. He observed the progress of the groups and visited particular groups as needed. In these visits he checked solutions, made corrections, gave hints, clarified notation, provided encouragement, and tried to help the groups function more smoothly.

The teacher used a democratic style of leadership [10] which involved considerable respect and friendliness toward the students. He did not give orders or disrupting commands. Instead, he offered guiding suggestions when the students wanted or clearly needed them. He used a minimal amount of constructive praise and criticism, usually directed to a group as a whole. Basically, he helped students to learn, rather than forcing them to learn.

Interest in the mathematical discussions was to be the major motivation; this required an increase in student freedom and a reduction in pressure. The students were free to explore mathematical questions that arose in their groups. The students decided whom to work with and when to change groups. The teacher used an A-B grading scale, and the students discussed grading policies and voted for take-home exams.

The discovery class met five periods per week for two semesters. The twelve students were all volunteers with A or B grades in high school mathematics and at least a mild interest in that subject. The students performed well on seven take-home examinations. In a final examination on basic facts and skills, the discovery class performed slightly better than a control class taught by the lecture-discussion system. However, the difference was not statistically significant, and it might have resulted from the special entrance requirements for the discovery class.

The students in the discovery class responded to an open-ended questionnaire, with the following results. On the negative side, most students were concerned for varying periods of time about covering enough material. Students sometimes became frustrated or angry, particularly when the mathematical problems were too hard. The students had difficulty at first in forming effective working groups. On the positive side, the pilot class had positive or null effects on each student's interest in mathematics and estimate of his problem solving skill. Almost all of the students had a closer, more personal relationship with their mathematics teacher than with their other teachers. Most students found their calculus class more stimulating than their other classes, and everyone's attitude toward the class either stayed the same or improved during the year.

The students' attitudes can be conveyed more vividly by quoting some questionnaire responses; no student is quoted more than once. (1) "Other students, no matter who, force you to learn more." (2) "Most classes stress being able to

use formulas while this stresses total understanding." (3) "It is my most interesting and liked class. I enjoy coming to it." (4) "I think I learned a lot more this year than I did in all three years of high school math." (5) "It showed me that I can do things that before looked impossible. All it takes is a little understanding. Math doesn't scare me as much now." (6) "I simply feel it was a great experiment (and experience) and more subjects should be adapted to this general method." (7) "This type of class was, in my estimation, the closest possible setup to an ideal learning situation."

After the pilot study, the author made three changes in his small group classes. First, the course grade was based largely on homework. The teacher checked some problems, and class members took turns checking the others [6]. Secondly, the teacher introduced new concepts and problems in written form, rather than in class discussions. Dittoed work sheets allowed each group to set its own pace. Finally, the teacher held a presession with one small group before planning each class meeting. He could then design work sheets that were interesting, challenging, and reasonable for students.

The author and other teachers have used the small group-discovery method in honors calculus, abstract algebra, and Euclidean geometry. These classes had roughly twenty-five students apiece. Conceivably, a small group approach with easier problems might be a realistic way to handle large enrollments without mass lectures. Such an approach would entail a special textbook, limited teacher guidance, and problems of suitable difficulty for the intended population.

References

1. Kenneth Cummins, A student experience-discovery approach to the teaching of calculus, The Mathematics Teacher, 53 (1960) 162–170.

2. N. A. Davidson, The Small Group-Discovery Method of Mathematics Instruction as Applied in Calculus (doctoral dissertation, University of Wisconsin, Madison, 1970), to appear as a Technical Report of the Wisconsin Research and Development Center for Cognitive Learning, Madison, 1971.

3. Morton Deutsch, The Effects of Cooperation and Competition upon Group Process, in D. Cartwright and A. Zander (eds.), Group Dynamics: Research and Theory, 2nd ed., Harper & Row, New York, 1960, pp. 414–448.

4. John Dewey, Democracy and Education, Macmillan, New York, 1916 (republished: Free Press Paperback ed., New York, 1966).

5. ———, Experience and Education, Kappa Delta Pi, New York, 1938 (republished: Collier Books Paperback ed., New York, 1963).

6. Donald Kingsbury, An experiment in education, unpublished manuscript available from the Mathematics Department, McGill University, Montreal 2, P. Q., Canada, 1963.

7. C. M. Larsen, The heuristic standpoint in the teaching of elementary calculus (doctoral dissertation, Stanford University, 1960), Dissertation Abstracts, 21, No. 9 (1961), 2632, 2633.

8. George Polya, Mathematical Discovery, vol. 2, Wiley, New York, 1965.

9. V. D. Turner, C. D. Alders, F. Hatfield, Harvey Croy, and Charles Sigrist, A study of ways of handling large classes in freshman mathematics, this MONTHLY, 73 (1966) 768–770.

10. Ralph White and Ronald Lippitt, Leader Behavior and Member Reaction in Three Social Climates, in D. Cartwright and A. Zander (eds.), Group Dynamics: Research and Theory. Harper & Row, New York, 1960, pp. 527–553.

$$\sum_{1 \leq i < j \leq n} \frac{|A_i \cap A_j|}{|A_i| \cdot |A_j|} < 1,$$

then the sets $A_1, \cdots, A_n$ have a system of distinct representatives (i.e., there are $a_1, a_2, \cdots, a_n$ such that $a_i \in A_i$ and $a_i \neq a_j$ for $i \neq j$).

E 2310. *Proposed by Hal Forsey, San Francisco State College*

Does there exist a positive function f such that if x is irrational and y is rational, then $f(x)f(y) \leqq |x-y|$?

E 2311. *Proposed by Huseyin Demir, Middle East Technical University, Ankara, Turkey*

Prove that, if a quadrilateral $A_1A_2A_3A_4$ can be inscribed in a circle, then the (six) lines drawn from the midpoints of A_pA_q perpendicular to A_rA_s (p, q, r, s are distinct) are concurrent.

E 2312. *Proposed by Huseyin Demir, Middle East Technical University, Ankara, Turkey*

Let D be a point in the plane of a positively oriented triangle ABC and let AD, BD, CD intersect the respective opposite sides in A_1, B_1, C_1. If the oriented segments $\overline{BA}_1$, $\overline{CB}_1$, $\overline{AC}_1$ are equal ($=\delta$), then D is uniquely determined and lies in the interior of ABC. (Notice the analogy between D and the Brocard point Ω.)

SOLUTIONS OF ELEMENTARY PROBLEMS

A Difficult Triangle Inequality

E 2245 [1970, 652]. *Proposed by A. W. Walker, Toronto, Canada*

If A, B, C; a, b, c; s are the angles, side lengths, and semi-perimeter of any plane triangle, then

$$(a+b+c)^3(s-a)(s-b)(s-c) \geqq (a^2+b^2+c^2)^3 \cos A \cos B \cos C.$$

Solution by the proposer. The result is obvious for right and for obtuse triangles and is easily verified for isosceles triangles. For all acute and right triangles we establish the stronger inequality

$$R^2(R-r+d)s^4 \geqq 4(r+d)(2R^2+r^2)^3, \tag{1}$$

with R, r, Δ denoting the circumradius, inradius, and area of triangle ABC, and $d \equiv (R^2-2\,Rr)^{1/2}$. It is known [1] that

$$4R^2 \cos A \cos B \cos C = s^2 - (2R+r)^2, \tag{2}$$

and also [2] that

$$4(2R^2+r^2) \geqq (a^2+b^2+c^2).$$

Hence the stated result will follow if we can show that

$$(Rs\Delta)^2 \geqq 2(2R^2 + r^2)^3[s^2 - (2R + r)^2]. \tag{3}$$

Blundon [3] has established the inequalities

$$2R^2 + 10Rr - r^2 + 2d(R - 2r) \geqq s^2 \geqq 2R^2 + 10Rr - r^2 - 2d(R - 2r). \tag{4}$$

From the first inequality in (4) we can show that

$$2r^2(r + d) \geqq (R - r + d)[s^2 - (2R + r)^2],$$

and (3) will therefore follow if we can prove (1).

The inequality (1) is a consequence of the following two results:

$$R^2(R - r + d)(2R + r)^4 \geqq 4(r + d)(2R^2 + r^2)^3, \tag{5a}$$

$$R^2(R - r + d)[2R^2 + 10Rr - r^2 - 2d(R - 2r)]^2 \geqq 4(r + d)(2R^2 + r^2)^3, \tag{5b}$$

in the intervals $0 \leqq r/R \leqq (\sqrt{2}-1)$ for (5a), and $(\sqrt{2}-1) \leqq r/R \leqq 1/2$ for (5b). (Neither (5a) nor (5b) is valid over the entire range $0 \leqq r/R \leqq 1/2$.) Note that from (2) the left side of (1) is $\geqq$ the left side of (5a) for all acute and right triangles, and from (4) the left side of (1) is $\geqq$ the left side of (5b).

To prove (5a) and (5b), we introduce a parameter x such that

$$\frac{d}{R} = \frac{x - 1}{x + 1}, \qquad \frac{r}{R} = \frac{2x}{(x + 1)^2}, \qquad (x \geqq 1),$$

and find, after tedious effort, that (5a) and (5b) are equivalent to

$$\text{(6a)}\; P_1(x) \equiv 2x^{13} + 17x^{12} + 52x^{11} + 29x^{10} - 226x^9 - 663x^8 - 816x^7 - 435x^6 + 10x^5 + 179x^4 + 128x^3 + 48x^2 + 10x + 1 \geqq 0, \qquad (x \geqq 1 + \sqrt{2});$$

$$\text{(6b)}\; P_2(x) \equiv -x^{14} - 14x^{13} - 31x^{12} + 48x^{11} + 253x^{10} + 270x^9 - 129x^8 - 484x^7 - 314x^6 + 34x^5 + 181x^4 + 128x^3 + 48x^2 + 10x + 1 \geqq 0, \qquad (1 \leqq x \leqq 1 + \sqrt{2}).$$

In $P_1(x)$ there are two variations in the signs of the coefficients, so by Descartes' rule, P_1 has at most *two* positive roots. But $P_1(0) > 0$, $P_1(1) < 0$, and $P_1(1+\sqrt{2}) > 0$, so that $P_1(x) > 0$ for $x \geqq 1 + \sqrt{2}$, and (5a) is established. Likewise P_2 has at most *three* positive roots; but $P_2(0) > 0$, $P_2(1) = P_2'(1) = 0$, $P_2(1+\sqrt{2}) > 0$, and $P_2(x) \to -\infty$ as $x \to \infty$. We see then that P_2 has a double root at $x = 1$ and that its only other positive root is greater than $1 + \sqrt{2}$; since $P_2(1+\sqrt{2}) > 0$ it follows that $P_2(x) \geqq 0$ for $1 \leqq x \leqq 1 + \sqrt{2}$ with equality if and only if $x = 1$. This establishes (5b). [One degenerate case deserves mention: if $r = 0$, then $x = \infty$ and the argument with P_1 does not apply. But if $r = 0$, then one side of triangle ABC is zero, and if $r = 0$, then (1) becomes simply $s \geqq 2R$ and this is trivially an equality. Hence it is shown that (1) holds in general, with equality if and only if $r/R = 0$ or $r/R = 1/2$, i.e., if and only if the triangle ABC is either degenerate or equilateral.—Ed.]

REMARK. Expressed in terms of a, b, c only, the given inequality takes the form

$$(*)\quad (abc)^2(a+b+c)^3(b+c-a)(c+a-b)(a+b-c) \geqq (a^2+b^2+c^2)^3(b^2+c^2-a^2)(c^2+a^2-b^2)(a^2+b^2-c^2).$$

This seems to be true for all nonnegative (a, b, c)—and has been verified by computer for about 1000 sets of (a, b, c) ranging from 10^{-6} to 10^3. The set $(4, 5, -6)$ shows that it is not true for all real numbers.

Two incorrect solutions were received.

References

1. This MONTHLY, 74 (1967) 566.
2. Elem. der Math., 18 (1963) 129.
3. Canadian Math. Bull., 8 (1965) 616.

Three Collinear Centroids—Three Noncollinear Circumcenters

E 2249 [1970, 765]. *Proposed by A. L. Holzhauer, Charlotte, N. C.*

Given any three triangles Δ_1, Δ_2, Δ_3. Let H_1, H_2, H_3 be the orthocenters; G_1, G_2, G_3 the centroids; and O_1, O_2, O_3 the circumcenters. Prove that the centroids of the triangles $H_1H_2H_3$, $G_1G_2G_3$ and $O_1O_2O_3$ are collinear, and their circumcenters are likewise collinear.

Solution by Simeon Reich, Israel Institute of Technology. We identify the points (upper case letters) with complex numbers (lower case letters). By Euler's theorem $g_i=(2o_i+h_i)/3$, $i=1, 2, 3$. Thus if the centroid of $H_1H_2H_3$ is $h_m=(h_1+h_2+h_3)/3$ and the centroid of $O_1O_2O_3$ is $o_m=(o_1+o_2+o_3)/3$, then the centroid of $G_1G_2G_3$ is $g_m=(g_1+g_2+g_3)/3=(2o_m+h_m)/3$. Hence the first result. Furthermore, we see that just as G is two-thirds of the way from H to O, so also is G_m two-thirds of the way from H_m to O_m.

As for the second stated result, note that, given any two points H, O (which may coincide), one can construct a triangle with orthocenter H and circumcenter O. Hence we may assume that in plane cartesian coordinates $H_1=(-1, 0)$, $H_2=(1, 0)$, $H_3=(0, 1)$, $O_1=(3, 0)$, $O_2=(1, 0)$, and $O_3=(2, 1)$. Then $G_1=(5/3, 0)$, $G_2=(1, 0)$, and $G_3=(4/3, 1)$. The circumcenter X of $H_1H_2H_3$ is $(0, 0)$ and that of $O_1O_2O_3$ is $Y=(2, 0)$. But Z, the circumcenter of $G_1G_2G_3$, cannot be collinear with X and Y since angle $G_1G_3G_2$ is not 90°. This refutes the second result.

Also solved by L. Carlitz & R. A. Scoville, C. S. Karuppan Chetty (India), Jordi Dou (Spain), A. J. Keeping, Peter Kornya, and Charles Wexler. Proofs of the first result only were submitted by M. Bolurizadeh (Iran), Michael Goldberg, and A. J. Papadopoulis (Greece).

A Dense Subset of the Reals

E 2250 [1970, 766]. *Proposed by C. A. Kottman, Louisiana State University*

Prove that, if one is given any rectangular sheet of paper and a number $\epsilon>0$, he may by repeated foldings of the paper in half or in thirds (lengthwise or widthwise or both) arrive at a smaller rectangle with ratio r of length to width satisfying $1-\epsilon\leqq r\leqq 1+\epsilon$.

Solution by G. A. Heuer, Concordia College. The conclusion follows from the fact that the numbers 2^j3^k, with integral j and k, are dense in the positive reals.

To see that this is so, recall that since $\log_3 2$ is irrational, the residues mod 1 of integral multiples of $\log_3 2$ are dense in (0, 1). (Cf. Theorem 6.3 in Niven, *Irrational Numbers*, Carus Monograph 11.) Thus, given $0 \leqq a < b$, there are integers j and k such that $\log_3 a < k + j \log_3 2 < \log_3 b$; i.e., that $a < 2^j 3^k < b$.

(The problem is a slight variation on E 1565 [1963, 1101]. See also Leo Moser and Nathaniel Macon, *On the distribution of first digits of powers*, Scripta Mathematica, XVI (1950) 290–292.)

Also solved by Anders Bager (Denmark), Jordi Dou (Spain), Neal Felsinger, Michael Goldberg, M. G. Greening (Australia), Ellen Hertz, Dean Hickerson, Peter Kornya, O. P. Lossers (Netherlands), Norman Miller, S. R. Murfry, P. H. Young and the proposer.

Lattice Points in Color

E 2251 [1970, 766]. *Proposed by T. C. Brown, Simon Fraser University, Burnaby, British Columbia*

Consider a rectangular array of dots with an even number of rows and an even number of columns. Color the dots either red or blue in such a way that every row has the same number of red and blue dots, and likewise every column. Whenever two dots of the same color are adjacent in the same row or column, connect them with a line segment of that color. Show that the total number of blue segments equals the total number of red segments.

Solution by S. B. Maurer, Phillips Exeter Academy. It suffices to prove that, for any two adjaçent columns (rows), the number of red lines bridging these columns (rows) equals the number of blue lines. Let us call a dot in either column *matched* if it is of the same color as the dot beside it in the other column; otherwise call it *mismatched*. Thus a dot is the endpoint of a line between the columns only if it is matched. Now, since both columns have the same number of red dots and the same number of matched red dots, they also have the same number of mismatched red dots; but clearly the number of mismatched red dots in one column is the number of mismatched blue dots in the other column. Hence, in either column, the number of mismatched red dots equals the number of mismatched blue dots. Finally, since in each column the total number of red dots equals the total number of blue dots, we see that the number of matched red dots equals the number of matched blue dots.

Also solved, in essentially the same way, by Gaile Fleming, Peter Kornya, Joel Levy, Norman Miller, Simon Reich (Israel), J. P. Robertson, G. F. Schumm, J. B. Wilker, and the proposer.

Solutions involving construction of matrices by assigning real numbers to dots according to whether they are red or blue were offered by D. J. Bordelon, E. J. Cockayne, M. J. Frank, A. J. Keeping, Harry Lass, F. D. Parker, and Gabriel Rosenberg.

Various other solutions were given by J. K. Bidwell, T. A. Brown, Mannis Charosh, Neal Felsinger, R. A. Gibbs, Michael Goldberg, M. G. Greening (Australia), Ellen Hertz, Dean Hickerson, Y. J. Inn, O. P. Lossers (Netherlands), J. P. McLean, Steven Morfey, and David Spear.

W. D. Bousma and Jordi Dou (Spain) showed that, in general, the difference between the number of red and blue segments depends on the difference between the totals of red and blue dots, and the differences in the border rows and columns.

An Easy Urn Problem

E2252 [1970, 766]. *Proposed by Harry Lass, Jet Propulsion Laboratory, California Institute of Technology*

Given n urns numbered $1, \cdots, n$ and k objects, with $k \leqq n$. Suppose each of the objects is placed at random in one of the urns. Find for $r=1, 2, \cdots, n$ the probability that the number of objects in the first r urns is less than or equal to r.

Solution by Richard Post, San Jose State College. We can order the objects and place them in the urns in k independent trials. If placing an object in the first r urns is considered a success, then the problem is equivalent to that of finding the probability of r or fewer successes in k Bernoulli trials where the probability of success is equal to r/n. It is well known that this is $n^{-k}\sum_{i=0}^{r}\binom{k}{i}r^i(n-r)^{k-i}$. (Note that the formula is valid even if $r>k$, since by the usual conventions for binomial coefficients, the sum is simply 1.)

Also solved by Ed Bednar, W. D. Bouwsma, R. L. Cramer, Jordi Dou (Spain), Neal Felsinger, Ellen Hertz, G. A. Heuer, J. C. Hickman, Thomas Hughes, N. J. Kuenzi, S. B. Leonard, Lonnie Machen, R. M. Meyer, Simeon Reich (Israel), Perry Scheinok, F. G. Schmitt, Jr., C. E. Skinner, C. L. Smith, David Spear, R. K. Tamaki, and E. H. Voorhees, Jr.

The Return of the Catalan Numbers

E 2253 [1970, 882]. *Proposed by L. A. Gehami, Glastonbury, Connecticut*

Find the number S_n of distinct n-tuples $(a_1, a_2, \cdots, a_n)$ such that a_i is a nonnegative integer, $i=1, 2, \cdots, n$; $\sum_{i=1}^{n} a_i=n$; $\sum_{i=1}^{k} a_i \geqq k$, $k=1, 2, \cdots, n$.

Solution by Robert Fray, Florida State University. We can associate with each n-tuple $(a_1, \cdots, a_n)$ the lattice path from $(0, 0)$ to (n, n) by connecting the points $(0,0)$, $(0, s_1)$, $(1, s_1)$, $(1, s_2)$, $(2, s_2)$, $(2, s_3)$, $\cdots$, $(n-1, s_n)$, (n, s_n), where s_k is the partial sum $s_k=a_1+a_2+\cdots+a_k$. This association sets up a one-to-one correspondence between the designated set of n-tuples and the set of lattice paths from $(0, 0)$ to (n, n) which never fall below the line $y=x$. It is well known (cf. William Feller, *An Introduction to Probability Theory and Its Applications*, 2nd edition, New York, 1957, p. 71) that the number of such paths is the nth Catalan number

$$C_n = \frac{1}{n+1}\binom{2n}{n}.$$

Also solved by M. T. Bird, L. Carlitz, Jordi Dou (Spain), David Fried & Robert Tax, C. S. Gardner, M. G. Greening (Australia), M. Hirschhorn, Thomas Hughes, R. M. Krause, Harry Lass, Eric Langford, C. L. Moffitt & Peter Geyer, Robert Patenaude, B. H. Rodin, D. P. Roselle, R. W. Sielaff, David Spear, R. Z. Vause (Saudi Arabia), M. R. Wise, and the proposer.

Editorial Note. Because of the many applications of the Catalan numbers, it is not surprising that there was a wide variety of solutions submitted. Some solvers set up a recurrence relation and solved it by induction, whereas others derived the generating function $f(x)=\sum S_n x^n$ and showed that $xf^2=f-1$. Others interpreted it as a ballot problem or as an enumeration of non-associative products. Problem E2054 [1969, 192] is quite similar.

ADVANCED PROBLEMS

All solutions of Advanced Problems should be sent to J. Barlaz, Rutgers—The State University, New Brunswick, N.J. 08903. Solutions of Advanced Problems in this issue should be typed (with double spacing) on separate, signed sheets and should be mailed before December 31, 1971. Contributors (in the United States) who desire acknowledgment of receipt of their solutions are asked to enclose self-addressed stamped postcards.

An asterisk (*) means neither the proposer nor the editors supplied a solution.

5808.* *Proposed by L.-S. Hahn, University of New Mexico*

A weak version of van der Waerden's theorem reads as follows: Let the set of all *natural numbers* be divided in any manner whatsoever into finitely many sets. Then an arithmetic progression of arbitrarily many terms can be found in at least one of these sets.

Now, known proofs of van der Waerden's theorem are elementary but highly non-simple. By replacing the set of all natural numbers by the set of all real numbers (or rational numbers, or even complex numbers) can one find a simple direct proof of the corresponding statement?

5809. *Proposed by Richard Stanley, Massachusetts Institute of Technology*

Let X be a topological space such that an arbitrary intersection of open sets is open. Show that if X is connected, compact and normal (without assuming T_0, T_1, or Hausdorff), then X is contractible.

5810. *Proposed by Simeon Reich, Israel Institute of Technology, Haifa*

Let $f(x)$ be continuous on $[a, b]$ and differentiable at a and at b. If $f'(a) = f'(b)$, then there is a number $H>0$ such that corresponding to any h, $0<h\leqq H$, there exists d, $a<d<b-h$, such that $[f(d+h)-f(d)]/h=[f(d)-f(a)]/(d-a)$.

5811*. *Proposed by T. C. Brown, Simon Fraser University, Burnaby, Canada*

Let S be a nonempty subset of the plane such that for each x in S exactly one of $x+(0, 1)$ and $x+(1, 0)$ also belongs to S. Prove or disprove that for each positive integer k there is a line in the plane (perhaps different lines for different k) which contains at least k points of S.

5812. *Proposed by Paul Monsky, Brandeis University*

If $f=x^4+4x^3-6x^2+4x+1$, evaluate

$$I = \int \frac{xdx}{\sqrt{f}} \cdot$$

5813. *Proposed by A. R. Barron, Brandeis University*

Show that if A is a bounded operator and B a self adjoint operator, on some Hilbert space, then

$$[B, [B, A]] = 0 \qquad \text{implies } [B, A] = 0.$$

Note: $[X, Y]=XY-YX$.

SOLUTIONS OF ADVANCED PROBLEMS

The Smallest Zero of a Polynomial

5741 [1970, 665]. *Proposed by Simeon Reich, Israel Institute of Technology*

Let $p(z)=a_0+a_1z+\cdots+a_nz^n$, $a_n\neq 0$, be a polynomial with complex coefficients. It is known (Morris Marden, *The Geometry of the Zeros*, AMS, New York, 1949, p. 98) that if z_1 is the zero of largest modulus, then

$$|z_1| < \left\{1+\sum_{j=0}^{n-1}|a_j/a_n|^2\right\}^{1/2}.$$

Prove that if z_n is the zero of smallest modulus, then

$$|z_n| < \left\{1+\frac{1}{n}\left(\sum_{j=0}^{n-1}|a_j/a_n|^2\right)\right\}^{1/2}.$$

Solution by Emeric Deutsch, Polytechnic Institute of Brooklyn.
Denoting the zeros of the polynomial by $z_1, z_2, \cdots, z_n$, with z_n the zero of smallest modulus, we have

$$|z_n|^n \leqq |z_1z_2\cdots z_n| = |a_0/a_n|,$$

whence $|z_n| \leqq |a_0/a_n|^{1/n}$.

This inequality is stronger than the one of the proposal. Indeed, denoting $\sum_{j=0}^{n-1}|a_j/a_n|^2$ by b, we have

$$|a_0/a_n|^{1/n} = (|a_0/a_n|^2)^{1/2n} \leqq b^{1/2n} < \{(1+b/n)^n\}^{1/2n} = (1+b/n)^{1/2}.$$

Also solved by A. A. Jagers (Netherlands), G. Schmeisser (Germany), and the proposers.

Extending Subgroups in a Group

5742 [1970, 655]. *Proposed by Anne Penfold Street, University of Alberta*

Let G be a group and A a subgroup of G. Let $x\in G$, $x\notin A$. We say *x augments* A if $A_x\equiv A\cup(x, x^{-1})$ is also a subgroup of G. Suppose A is a subgroup of G such that any element of G augments A. Characterize G.

I. *Solution by C. V. Heuer and G. A. Heuer, Concordia College.* Every group satisfies this with $A=G$, so we suppose G has such a subgroup $A\neq G$.

For $x\notin A$, $|A_x|\leqq|A|+2$, where $|A|$ is the order of A. Since a proper subgroup of A_x can contain at most half of the elements of A_x, it follows that $|A|\leqq 2$.

If $|A|=1$, then every nonidentity element of G augments A if and only if every such element of G has order 2 or 3.

If $|A|=2$, then G is either the cyclic group of order 4 or the quaternion group. To see this let $x\in G$, $x\notin A$, and note that $x\neq x^{-1}$. Then A_x is cyclic of order 4. It follows that G contains exactly one element of order 2 and all other non-

identity elements must have order 4. One checks that the above mentioned groups are the only such groups having order 4 or 8. There are no such groups of order 16. (See, e.g., Hall and Senior, *The Groups of order* $2^n (n \leqq 6)$.) It follows that there are no such finite groups of order >16 since they would have to contain a subgroup of order 16 with the same property. Finally, there are no infinite groups with this property since every finitely generated subgroup of such a group would be finite (being of exponent 4) and many of these would have order $\geqq 16$.

II. *Addendum by the proposer.* The groups in which every element has order 2 or 3 have been completely determined by B. H. Neumann (*Groups whose elements have bounded orders*, J. London Math. Soc. 1, 12 (1937), 195–198) and there are four additional possibilities for G: (1) A direct product of groups of order 2; (2) A direct product of noncyclic groups of order 4, extended by an automorphism of order 3 which induces an outer automorphism in each of the subgroups of order 4; (3) A direct product of groups of order 3, extended by an automorphism of order 2 which transforms each element into its inverse; (4) A factor group of the Burnside group of exponent 3.

Also solved by T. L. Bartlow, E. D. Bolker, J. P. Celenza, Neal Felsinger, S. H. Friedberg, D. A. Hejhal, John Israel, A. A. Jagers (Netherlands), Gustav Lehrer (Norway), R. C. Lyndon, N. S. Natarajan (India), H. Niederreiter, Roy Olson, E. F. Schmeichel, and A. M. Vaidya & V. S. Joshi (India).

As pointed out by the proposer and others, the cyclic nature of A_x (order 2, 3, or 4) is the substance of problem E 1657 [1964, 1134]. Lyndon gives a characterization of groups in which each element has order 2 or p (an odd prime) similar to that above.

Order of Derivatives of Real Functions in D_k

5743 [1970, 655]. *Proposed by Bjarni Jonssen and J. B. Nation, Vanderbilt University*

Let $f(x)$ be a real valued function with at least k derivatives. Given that for some real number r,

$$\lim_{x\to\infty} x^r f(x) = 0 \quad \text{and} \quad \lim_{x\to\infty} x^r f^{(k)}(x) = 0,$$

prove that $\lim_{x\to\infty} x^r f^{(j)}(x) = 0$, $0 \leqq j \leqq k$.

I. *Solution by L. S. Bosanquet, University of Western Ontario.* If $k \geqq 2$ and $\Delta f(x) = f(x+1) - f(x)$, then

$$\begin{aligned}\Delta^{k-1} f(x) &= \int_0^1 dt_1 \cdots \int_0^1 dt_{k-2} \int_0^1 f^{(k-1)}(x + t_1 + \cdots + t_{k-1}) dt_{k-1} \\ &= f^{(k-1)}(x+\xi) \qquad (0 \leqq \xi \leqq k-1),\end{aligned}$$

since $f^{(k-1)}(x)$ is continuous. Hence

$$\Delta^{k-1} f(x) - f^{(k-1)}(x) = \xi f^{(k)}(x + \theta\xi) \qquad (0 < \theta < 1),$$

and so $o(x^{-r})-f^{(k-1)}(x)=o(x^{-r})$ as $x\to\infty$, i.e. $\lim_{x\to\infty} x^r f^{(k-1)}(x)=0$. The result follows inductively.

II. *Solution by Stan Rajnak, Kalamazoo College.* For each integer j, $1<j<k$, expand $f(x+j)$ in a Taylor polynomial about the point x:

$$f(x+j)=f(x)+jf'(x)+\frac{j^2f''(x)}{2!}+\cdots+\frac{j^{k-1}f^{(k-1)}(x)}{(k-1)!}+\frac{j^kf^{(k)}(\theta_j)}{k!},$$

where $x<\theta_j<x+j$. This may be considered as a system of linear equations in the unknowns $f(x)$, $f'(x)$, $\cdots$, $f^{(k-1)}(x)$. The matrix of coefficients has for its ith row

$$1\ \frac{i}{1!}\ \frac{i^2}{2!}\cdots\frac{i^{k-1}}{(k-1)!}.$$

From the corresponding determinant we may factor out the denominators common to the elements of each column and will have $(1!2!\cdots(k-1)!)^{-1}$ times the familiar Vandermond determinant. Hence the determinant of coefficients equals 1. The system of equations has therefore a solution and this solution gives $f^{(j)}(x)$ as a linear combination of $f(x+1)$, $f(x+2)$, $\cdots$, $f(x+k)$, $f^{(k)}(\theta_1)$, $\cdots$, $f^{(k)}(\theta_k)$. Now if $x<\xi<x+k$, then

$$\lim_{x\to\infty} x^rf(\xi)=\lim_{x\to\infty}\left(\frac{x}{\xi}\right)^r\xi^rf(\xi)=\lim_{x\to\infty}\left(\frac{x}{\xi}\right)^r\lim_{\xi\to\infty}\xi^rf(\xi)=1\cdot 0=0,$$

and similarly $\lim_{x\to\infty} x^rf^{(k)}(\xi)=0$. Hence $x^rf^{(j)}(x)$ is a linear combination of terms of the form $x^rf(\xi)$ or $x^rf^{(k)}(\xi)$ where $x<\xi<x+k$. Since these all go to 0 as $x\to\infty$, the result follows.

Also solved by Roger Giudici, D. A. Hejhal, A. A. Jagers (Netherlands), Joel Levy, and the proposers.

Packing Cubes with Bricks

5744 [1970, 656]. *Proposed by Jan Mycielski, University of Colorado*

A cube $20\times20\times20$ is built out of bricks of the form $2\times2\times1$. The faces of the bricks are parallel to the faces of the cube but they need not all lie flat. Prove that the cube can be pierced by a straight line perpendicular to one of the faces which does not pierce any of the bricks.

Solution by Bill Sands, University of Manitoba. Divide the cube into unit cubes. This defines a 20×20 grid on each face. We may assume that any line perpendicular to one of the faces will intersect the face at a grid point. There are $19^2=361$ such lines intersecting each face, or $361\cdot3=1083$ distinct lines of the above type passing through the cube. Each $2\times2\times1$ block will be pierced by exactly one of these lines. Consider such a line. Construct the two planes containing this line and perpendicular to the faces of the cube; this divides the cube

into quadrants, each containing an even number of unit cubes. Choose one of these quadrants, call it A. Any $2\times 2\times 1$ block that intersects A will intersect it in one, two, or four unit cubes. The blocks that intersect A in one unit cube are precisely the blocks that are pierced by the above line. Since A contains an even number of unit cubes, there must be an even number of blocks that intersect it in one unit cube, and so there must be an even number of blocks that are pierced by this line. Now 2166 blocks are required if every line is to pierce two bricks, whereas a $20\times 20\times 20$ cube will contain only 2000 blocks. Therefore there is at least one (in fact at least 83) lines that do not pierce any of the blocks.

Also solved by W. E. Hosken, Robert Israel, A. A. Jagers (Netherlands), D. A. Klarner & R. P. Nederpelt (Netherlands), R. M. Robinson, David Singmaster (England), Dean E. Smith, and the proposer.

Note. As observed by most of the solvers, the method works for any cube with edge not exceeding 20. But the proposer leaves us with the query of what happens with a cube having edge 22.

Covering Systems of Congruences

5747 [1970, 775]. *Proposed by H. M. Edgar and Martin Billik, San Jose State College*

Let $1<n_1<n_2<\cdots<n_k$ be integers. Let the integers b_i satisfy $0\leqq b_i\leqq n_i-1$ for every value of i with $1\leqq i\leqq k$, and assume that (n_i, n_j) is not a divisor of (b_i-b_j) for all $i\neq j$, $1\leqq i,j\leqq k$. Prove that there must exist an integer x satisfying $x\not\equiv b_i \pmod{n_i}$ for all i with $1\leqq i\leqq k$.

Solution by T. R. Butts, Michigan State University. First we rephrase the problem. The set $\{(b_j, n_j)\}_{j=1}^k$ is called a *covering* if every integer satisfies at least one of the congruences $x\equiv b_j \pmod{n_j}$; $j=1, 2, \cdots, k$. If every integer satisfies exactly one of the congruences, the covering is called *disjoint.*

Now the condition $(n_i, n_j)\nmid(b_i-b_j)$ for $i\neq j$ means that x cannot satisfy simultaneously two of the congruences and so if $1<n_1<n_2<\cdots<n_k$, it is enough to prove that there is no disjoint covering $\{(b_j, n_j)\}_{j=1}^k$.

Proof. Let $n=\text{l.c.m.}\{n_1, n_2, \cdots, n_k\}$ and define $f(x)=e^{2\pi ix/n}$ for x integral. Then if $x=b_j+ln_j$, it follows that

$$f(x) = e^{2\pi ib_j/n}\cdot e^{2\pi iln_j/n};$$

that is, $f(x)$ is a root of the equation

$$y^{n/n_j} = e^{2\pi ib_j/n_j}.$$

Thus the existence of a disjoint covering $\{(b_j, n_j)\}_{j=1}^k$ implies

$$y^n - 1 = \prod_{j=1}^{k}(y^{n/n_j} - e^{2\pi ib_j/n_j})$$

since both polynomials have the same roots. But the coefficient of y^{n/n_1} is zero, and since

$$\frac{n}{n_1} > \frac{n}{n_2} > \cdots > \frac{n}{n_k},$$

we must have

$$(-1)^{k-1}\prod_{j=1}^{k-1} e^{2\pi i b_j/n_j} = 0,$$

an obvious contradiction.

Other proofs of this result appear in the literature: See S. K. Stern, *Math. Annalen* 134 (1958), 289–294 and P. Erdös, *Matematikai Lapok* 3, No. 2 (1952), 122–128.

Also solved by Simeon Reich (Israel), J. M. Sherrill, and by the proposer.

A Divergence Test for Infinite Series

5748 [1970, 775]. *Proposed by R. L. Graham, Bell Telephone Laboratories, Murray Hill, N. J.*

Let $0 < a_n < a_{n+1} + a_n^2$, $n \geqq 1$. Show that $\sum_{n=1}^{\infty} a_n$ diverges.

Editorial Note. The solution is completely carried out in a note by K. A. Post, *A combinatorial lemma involving a divergence criterion for series of positive terms*, this MONTHLY 77 (1970), pp. 1085–1087. The editors offer apologies for the oversight which permitted this duplication.

Also solved by M. T. Bird, David Borwein, D. W. Boyd, J. Komlos (Hungary), D. E. Manes, Amram Meir, D. J. Newman, and Henry Ricardo.

REVIEWS

EDITED BY J. ARTHUR SEEBACH, JR. AND LYNN A. STEEN

with the assistance of the mathematics departments of St. Olaf and Carleton Colleges

COLLABORATING EDITOR FOR FILMS: SEYMOUR SCHUSTER, Carleton College

Printed materials for review should be sent to: Book Review Editor, American Mathematical Monthly, St. Olaf College, Northfield, MN 55057. Films and correspondence relating to films should be sent to Seymour Schuster, Carleton College, Northfield, MN 55057.

All unsigned material is written by the editors. A boldface capital **C** *in the margin indicates that a review is based in part on classroom use. Professors willing to write such a review should inform the editor in order to avoid duplication.*

Counterexamples in Topology. By Lynn A. Steen and J. Arthur Seebach, Jr. Holt, Rinehart, and Winston, New York, 1970. 210 pp. \$9.50. (Telegraphic Review, March 1971.)

Counterexamples in Topology is a valuable addition to the small collection of books which I keep on the shelf in my office. I think other teachers and students may be almost as subject to forgetting a definition of a topological property or the construction of a simple example as I am, and they too will enjoy having this book within arm's reach. The book is completely unique; no other book now in print serves its purpose.

Its purpose is very narrow. It seems a counterexample to the title that you will not find a Möbius band, Alexander's horned sphere, Antoine's necklace, a Klein bottle, or a dunce's cap described here. The examples are all from very general axiomatic point set topology, not touching set theory on one end or Euclidean topology on the other. But if you want to find a space which is T_0, T_4, T_5, compact, first countable, arc connected, 2nd category, but is not regular or metrizable and yet has a dispersion point, you need only look in the back of the book.

It should also be mentioned that the examples are all elementary. The authors used undergraduates to help them compile the examples, and few complicated or sophisticated examples are given. These are the simple examples which are part of everyone's heritage. Each example has been checked to see which of the defined properties it has—and it is this list which sounds formidable. The defined properties are also only the more common and widely used ones. In fact, considering the mass of definitions in this area, the authors were extremely conservative in their choices.

The book is well organized and easy to use; it reads almost like a Sears catalogue. I mean this as a compliment. There is no nonsense. The first 38 pages consist of definitions of properties from *topological space* to *pseudometrizable* with their relations to each other diagrammed and the words to be defined printed in large bold-faced type. The next 118 pages consist of examples, concisely and accurately defined and in some way also organized so they fit together in groups with each given a name to remember it by and a number for easier cross reference and looking up. Then there are 17 pages of cross references followed by a few problems and comments and a short bibliography.

Axiomatic topology, the area covered by this book, is a strange mathematical field for it has few deep theorems and general theories; it is a world of counterexamples. The absence of other books of examples in the area is surely based on the fact that the examples are so diverse and complicated that organizing them is nearly impossible. In other mathematical fields one restricts one's problem by requiring that the space be Hausdorff or paracompact or metric, and usually one doesn't really care which, so long as the restriction is strong enough to avoid this dense forest of counterexamples. A usable map of the forest is a fine thing. Students who sometimes find this maze impenetrable should be very happy to have this book as a guide. And even those of us who work exactly in the area will profit from its organization. It may even promote more uniformity in names used to describe common examples. This book is a needed welcome addition.

Everyone loves to find counterexamples. For instance, a single point space is a counterexample to the authors' statement on page 31 that no connected set can be totally disconnected.

The above review was solicited and edited by the previous editor of the Review Section.

MARY ELLEN RUDIN, University of Wisconsin

Matrix Algebra, A Programmed Introduction. By Richard C. Dorf. Wiley, New York, 1969. viii+260 pp. $7.95 ($5.95 paper). (Telegraphic Review, January 1970.)

This book is of no use to any reader of the MONTHLY. Except for the programmed format and 8 pages on the exponential function and differentiation of a matrix (and 8 pages is nothing—it takes 60 pages to describe matrices, their addition and multiplication), the tenor of the book is *fin de siècle.* It might be appropriate for eighth grade or retarded ninth grade students, but is completely inadequate even for a matrix course which is specifically designed for the author's announced audiences in fields of application.

The author is either careless or unknowing. For example, he describes a determinant as a number and then discusses the rows and columns of a determinant. For real t, exp t is described as a polynomial function. In addition to other such errors, there is a frequent tone of condescension.

P. H. YEAROUT, Brigham Young University

Elementary Number Theory. By Underwood Dudley. Freeman, San Francisco, 1969. ix+262 pages. $8.50. (Telegraphic Review, January 1970.)

This book is distinguished mainly by the lucid and lively style of the author. With few exceptions, the statements of the theorems are carefully worded to convey information as easily as possible. For example, the quadratic reciprocity theorem is stated: $(p/q)=(q/p)$, unless $p\equiv q\equiv 3 \pmod 4$ in which case $(p/q) = -(q/p)$. (This is the way it should be stated, but most authors prefer the arcane formula $(p/q)(q/p)=(-1)^{(p-1)/2\cdot(q-1)/2}$.) Likewise, proofs are carefully done, and examples are nicely woven into both statements and proofs to bring the reader to a quicker understanding. However, there are a few lapses which could be profitably eliminated.

The topics covered are standard—congruences, quadratic reciprocity, decimal expansions of rationals, simple diophantine equations, and a weak form of the prime number theorem. The book is suitable for undergraduates, or even good high school students, though not necessarily for the very best junior or senior math majors. There is a tremendous number of exercises of varying difficulty, with sections of hints and answers, and many stimulating but unpretentious historical remarks are scattered throughout.

My only reservation is that there are some annoying errors and the occasional misleading statement. Except for these slips, it is highly suitable for independent study. The section on decimal expansions has several errors on pages 119 and 120, and there is confusion between integers and positive integers in section 16 on $x^2+y^2=z^2$. Also the running hypotheses on the integers are never completely spelled out (which is reasonable) but at the beginning they vary a bit without warning. The author has prepared a list of errata which is available on request.

CARL RIEHM, University of Notre Dame

C *Théorie Algébrique des Nombres*. By Pierre Samuel. Hermann, Paris, 1967. 130 pp. 18F. (Telegraphic Review of English translation, April 1971.)

This small book of less than 100 pages of actual text was used at the University of Toronto (Spring term, 1970) for a fourth year course in algebraic number theory—in addition to fourth year students, there were third year students as well as graduates. The stronger third year students found the treatment by Samuel quite comprehensible; I might add that they were concurrently taking their first course in modern algebra (at the Lang level), having had previously a good grounding in linear algebra. Although the book is in French, the students found it easy to read with a small French vocabulary. The only demand that the author makes on the student is that he have the mathematical maturity to *understand* Galois Theory. A nice feature of this book is that it is self-contained; indeed, a modest amount of Noetherian ring and module theory is developed, and even a proof of the Fundamental Theorem of Galois Theory over fields of characteristic zero is to be found.

Within the short span of twenty-six hours of lectures, I was able to give complete proofs (following Samuel) of the following results, leaving very little detail to be supplied by the student: the finiteness of the class number and Dirichlet's Theorem on Units; Dedekind's Discriminant Theorem on ramification of prime ideals; Kummer's Decomposition Theorem (unfortunately, this does not appear in Samuel's book, except in a special case in section 5.4; however, a convenient version appears at the end of the first volume of Zariski and Samuel's *Commutative Algebra*); and Galois Extensions. The book ends with a beautiful application of Galois Theory to give a proof of Gauss' Quadratic Reciprocity Law via the Frobenius automorphism, together with earlier results on how prime ideals in $\mathbf{Z}$ split up under a quadratic extension of $\mathbf{Q}$. Needless to say, it is necessary to plan such a course carefully in order to fit all this into twenty-six hours, and still have several hours left for examples and some applications. In fact, several hours were saved by avoiding the material on localization, which meant that I could only prove Dedekind's Discriminant Theorem over a principal Dedekind domain, but this suffices in a first course.

The book contains twelve pages of exercises, ranging in difficulty from trivial to challenging. The book has two shortcomings. First, there are many printing errors and inaccuracies, even in the exercises. However, when errors appear in exercises, it considerably increases their value! Also, on a few occasions, shorter proofs can be found than those suggested by Samuel. Secondly, with *trivial* modifications, much of the material carries over to finite separable extensions, thus allowing function fields of one variable over a finite field as further examples of many of the theorems, and thus opening the door to many other beautiful theorems, e.g., the Riemann-Roch Theorem in characteristic p.

In conclusion, this book is to be highly recommended as an introduction to algebraic number theory. It was written with the student in mind, and rapidly introduces him to a fascinating branch of mathematics.

R. A. Smith, University of Toronto

Elementary Geometry. By Vincent H. Haag, Clarence E. Hardgrove, and Shirley A. Hill. Addison-Wesley, Reading, 1970. 266 pp. $8.95.

This textbook contains sufficient material on synthetic and analytic euclidean geometry to adequately prepare a K-6 teacher. (The teacher would also need a course in sets and rational arithmetic, with a few words about the reals, to complete her (his) minimal mathematical training.)

The material is presented in about the same order in which it is given in school. Chapter 1, "Informal Geometry," provides good physical motivations for basic geometrical concepts. The copious drawings are generally useful, although the method described for comparing angles is awkward if the angles are acute. Euler's simplex theorem is illustrated in an exercise. Chapter 2, "Measurement of Geometric Figures," is a rather careful discussion of length, area, volume, and angle. The mention of Wallis' formula and the discussion of real numbers and nested intervals should probably be skipped. Chapter 3, "Deductive Geometry and Constructions," introduces postulates of incidence, measure and congruence, which are used to justify some constructions and to prove a few theorems. But the main concern is with the nature of proof in mathematics. Chapter 4 introduces basic function concepts and rigid motions. Chapter 5, "Coordinates and Vector Geometry," introduces cartesian coordinates and vectors by means of translations. Chapter 6, "Geometric Transformations," is billed as " · · · a challenge to the reader who plans further study in mathematics · · · ." It amounts to an introduction to linear algebra, including orthogonal transformations and dilations.

The authors seem to be well aware of the lack of mathematical sophistication of their typical readers. Many attempts are made to reach the reader through her (his) interest in young children. It would be important to spend a lot of class time on the various "Teaching Questions and Projects." If this is done, the book should be quite successful.

R. J. BUMCROT, Hofstra University

C *Introduction to Complex Analysis.* By Rolf Nevanlinna and V. Paatero, translated by T. Kövari and G. S. Goodman. Addison-Wesley, Reading, Mass., 1969. ix+348 pp. $11.50.

This book uses the approach of the Finnish school—principally geometric and intuitive. However, that does not mean it is an easy book to use; some mathematical maturity seems quite necessary in order to read it.

When the reviewer used it as the text for an undergraduate course, he found the going very heavy. Most of the students had difficulty reading it. Although only analytic geometry and calculus are assumed, the instructor must fill in quite a bit. For instance, only slightly more than a page is devoted to the topology of the complex plane, and Mittag-Leffler's theorem on meromorphic functions is left as an exercise.

This book has many good features. The style is informal and very pleasant. A

collection of over 300 exercises is included; most of the exercises are very good, and the reader will benefit greatly in attempting them. The reviewer especially enjoyed reading the chapters on elementary functions which, as stated by the authors, followed the presentation given by E. Lindelöf.

In conclusion, the reviewer feels that this book should be possessed by every serious student of mathematics, whether his field is analysis or not. Nevertheless, as a text book, it may only be suitable in a course for honors or graduate students.

H. S. SUN, Fresno State College

Algebra. By Jacob K. Goldhaber and Gertrude Ehrlich. Macmillan, New York, 1970. 418 pp. $11.95. (Telegraphic Review, April 1970.)

A new one-volume textbook for a first graduate algebra course is a welcome addition to the literature, if only because there are so few available. Instructors who have had difficulty teaching out of Lang's *Algebra* will want to try this one. I think this book is a little less difficult than Lang, but an undergraduate course in algebra is definitely a prerequisite.

The main strength of this book is its exercises. There are altogether more than 450 problems of varying difficulty, which are collected at the end of each chapter. They will be a great help to the instructor, but he should be warned that the concrete examples which make algebra interesting are, generally speaking, in the exercises and not in the text.

As to content, the book has a fairly standard collection of topics in groups, rings, modules, and fields. It contains many "categorical ideas" and a little category theory. There are chapters on rank-one valuations (but, surprisingly, power series fields are omitted), noetherian rings and Dedekind domains, and the Wedderburn-Artin theory of semisimple rings. Group representations are omitted, but the greatest disappointment is the lack of any mention of ordered and real-closed fields. In spite of these omissions it is easy to believe the authors' statement that the book contains enough material for three leisurely semesters.

The book tends to move slowly from the very general to the specific. For example, in the field theory chapter algebraic and transcendental field extensions are treated together whenever possible, and the Galois theory of finite extensions comes 40 pages after the beginning of the chapter. I think that this style increases the importance of the exercises and of concrete examples in lecture. On the other hand, a student who gets through this chapter without bogging down will have a thorough working knowledge of field theory and a good appreciation of the unity of a set of algebraic ideas.

Finally, although I think this book has a good chance to be a fine text, I would not recommend it as a reference. The reason is mainly the lack of some basic topics. In addition I had trouble picking out isolated results from the later chapters.

B. F. WYMAN, University of Oslo and Stanford University

TELEGRAPHIC REVIEWS

Telegraphic reviews are intended to give prompt notice of books, with sufficient information to assist in deciding whether to order an examination copy or suggest library purchase. Possible uses are indicated as follows:

B = college bookstore stock L = library purchase
P = professional reading S = supplementary reading
T = textbook E = teacher education
13 to 18 = freshman to second year graduate level usage
1 to 4 = approximate time in semesters to cover text
* = positive emphasis ? = negative emphasis

Books on high-school material (pre-calculus) are denoted REMEDIAL, and normally receive telegraphic reviews only if they are written for college students. Publishers are denoted by the standard abbreviations used in *Books in Print*, which gives complete addresses.

ALGEBRA, T(14: 1). *Linear Algebra.* Henry G. Jacob and Duane W. Bailey. Houghton-Mifflin, 1971, xviii + 462 pp, $9.95. A text designed to follow a first year of the calculus. The approach is to give a short introductory chapter on geometric vectors and then to proceed with vector spaces, linear equations, inner product spaces, transformations, isometries and finally linear and bilinear forms. Examples and problems make good use of the applications of linear algebra to its many related areas in order to leave the student with an understanding of the place of linear algebra in the context of mathematics as a whole. L.L.K.

ALGEBRA, P, L. *Lecture Notes in Mathematics-149: K-Theory of Finite Groups and Orders.* Richard G. Swan and E. Graham Evans. Notes: E. Graham Evans. Springer-Verlag, 1970, 237 pp, $5.80 (P). A continuation of Swan's notes in this series (1968). These notes were used for an advanced seminar. W.C.R.

ALGEBRA AND NUMBER THEORY, POLYNOMIALS (OVER FINITE FIELDS) WITH GAPS, P, L(RESEARCH). *Lückenhafte Polynome über endlichen Körpern.* Lászlo Rédei. Birkhauser Verlag, 1970, 271 pp, $10. A treatise on the determination of the polynomials over finite fields which have certain gaps and which factor into linear factors over the coefficient field. Related to the work of Hajós on the factorization of finite abelian groups into complexes. J.D.-B.

ALGEBRA AND NUMBER THEORY, P, L. *Lecture Notes in Mathematics-177: Torsion Theories, Additive Semantics, and Rings of Quotients.* Joachim Lambek. Springer-Verlag, 1971, vii + 94 pp, $3.50 (P). A torsion theory on a category of R-modules specifies two classes of modules designated "torsion" and "torsion free." Specializations include a proof of the Freyd-Mitchell Embedding theorem and the construction of a generalized ring of quotients. An appendix by H. Storrer relates the concepts of torsion theory and dominant dimension. L.A.S.

ALGEBRA AND NUMBER THEORY, P, L. *Introduction to Affine Algebraic Groups.* G. Hochschild. Holden-Day, 1971, vii + 116 pp, $10.50. "Oriented toward representation theory, ...its principal aim is to fuse representation-theoretical technique with the elementary theory

of affine algebraic groups so as to provide an efficient tool, especially for Lie group theory." L.A.S.

ALGEBRA AND NUMBER THEORY, P, L, *Lecture Notes in Mathematics-156: Ample Subvarieties of Algebraic Varieties.* Robin Hartshorne. Springer-Verlag, 1970, xiv + 256 pp, $5.80 (P). Beginning with the equivalence of various characterizations of an ample divisor in codimension one, the author investigates the varied ways in which they generalize to higher codimension. The concluding chapter summarizes analogous analytic results. Huge bibliography! (From a course at the Tata Institute, 1969-70.) L.A.S.

ALGEBRA AND NUMBER THEORY, P, L, *Eléments de Géométrie Algébrique I.* A. Grothendieck and J.A. Dieudonné. *Die Grundlehren der mathematischen Wissenschaften in Einzeldarstellungen, Band 166.* Springer-Verlag, 1971, ix + 466 pp, $24.30. A revised version of Chapters 0 (Préliminaires) and Chapter I (Le language des schémas) which have previously appeared separately in several volumes in the IHES series. L.A.S.

ALGEBRA AND NUMBER THEORY, T(16-18: 1), S, P, L, *Topics in M-adic Topologies.* Silvio Greco and Paolo Salmon. *Ergebnisse der Mathematik und ihrer Grenzgebiete, Band 58.* Springer-Verlag, 1971, vi + 74 pp, $6.90. "The aim is to collect criteria concerning the ascent (from A to its M-adic completion Â) and descent (from Â to A) of several properties of commutative rings." This would be a good choice for a student seminar. L.A.S.

ALGEBRA AND NUMBER THEORY, T(18: 1, 2), P, L, *Lecture Notes in Mathematics-181: Separable Algebras Over Commutative Rings.* Frank DeMeyer and Edward Ingraham. Springer-Verlag, 1971, 157 pp, $4.60 (P). A presentation of the Brauer group and Galois theory for commutative rings. A historical section gives references to a long bibliography, and suggests problems still open. Some exercises. L.A.S.

ALGEBRA AND NUMBER THEORY, ANALYTIC NUMBER THEORY, P, L(RESEARCH). *Numbers with Small Prime Factors, and the Least κth Power Non-Residue. Memoirs of the American Mathematical Society, Number 106.* Karl K. Norton. AMS, 1971, 108 pp, $2.10 (P). Deals with the distribution of integers which have only relatively small prime factors and are terms of an arithmetic progression or relatively prime to a given number, with estimates of the least κth power non-residue to any modulus, with upper bounds for this nonresidue when the modulus is prime, and with a specific upper bound for the number of distinct prime factors of any integer. J.D.-B.

ALGEBRA, GALOIS THEORY, T(15-17: 1), S*, B, L, *Classical Galois Theory With Examples.* Lisl Gaal. Markham, 1971, x + 250 pp, $6.50 (P). Examples are central to this fill-in-the-blanks treatment of the fundamental theorem of Galois theory and the solvability criterion. It would be excellent for undergraduate independent study since it demands reader participation, or as a supplement to standard lecture treatments. The chapter on applications includes an "algorithmic" method for solving a solvable equation by first finding the Galois group. (Not only do we have to put up with ads on the back cover, but the publisher has the blurbs mismatched to

the titles.) L.A.S.

ALGEBRA, GROUPS, P*, B, L. *Lecture Notes on Nilpotent Groups, Number 2.* Gilbert Baumslag. AMS, 1971, viii + 73 pp, $3.10 (P). Finitely generated nilpotent groups are discussed in 10 expository lectures which were given at the University of Texas in 1969. An excellent and extensive bibliography is included. Perhaps an enjoyable book for a college teacher to go through. W.C.R.

ANALYSIS, T(15-16: 2). *Real Analysis.* Norman B. Haaser and Joseph A. Sullivan. Van Nostrand, 1971, ix + 341 pp, $11.95. This is a text for a first course in abstract analysis. After three preliminary chapters on sets, real number system and linear spaces, the main topics are covered: metric spaces, Lebesgue integral, normed linear spaces, Stieltjes integrals, and a chapter on inner product spaces and orthogonal bases. Should reinforce and deepen the student's understanding of the basic concepts on analysis without being too imposing. L.L.K.

ANALYSIS, T(18), P, L. *Holomorphic Functions, Domains of Holomorphy and Local Properties. Mathematics Studies, #1.* Leopoldo Nachbin. North Holland, 1970, vii + 122 pp, $4.95 (P). Based on a series of lectures at the University of Rochester, the book is intended as an elementary introduction to several complex variables, with an emphasis on how the theory differs from the theory of one complex variable. To this end, domains of holomorphy and local properties are studied in greater detail. Contains very little descriptive material. T.A.V.

ANALYSIS, INEQUALITIES, P. *Inequalities--II: Proceedings of the Second Symposium on Inequalities.* Ed: Oved Shisha. Acad Pr, 1970, xvi + 439 pp. A wide variety of research papers based on one and two-hour lectures at the Second Symposium on Inequalities held at the U.S. Air Force Academy, Colorado, in August 1967. Some have appeared in Journal of Approximation Theory, Volume 2 (1969). Contributions by N. Aronszajn, E.F. Beckenbach, G.T. Cargo, Ky Fan, Robert R. Kallman, John B. Kelly, H.W. McLaughlin, Marvin Marcus, F.T. Metcalf, B. Mond, T.S. Motzkin, Harry Pollard, Ray Redheffer, Paul C. Rosenbloom, Gian-Carlo Rota, Donald G. Saari, I.J. Schoenberg, O. Shisha, William Stenger, Olga Taussky, J.L. Walsh, Zvi Ziegler, and A. Zygmund. R.B.K.

APPLIED MATHEMATICS AND QUANTUM MECHANICS, T(17: 2), P, L. *Quantum Mechanics in Hilbert Space.* Eduard Prugovecki. Acad Pr, 1971, xv + 648 pp, $29.50. This book contains a wealth of information on the application of Hilbert space to non-relativistic quantum mechanics. The first half develops the mathematics of Hilbert space, measure theory, and operators in Hilbert space. The second half is devoted to the axiomatic structure of quantum mechanics and quantum mechanical scattering theory. Unfortunately the several axioms are scattered and only one can be located with the index. R.B.K.

APPLIED MATHEMATICS, CONTROL THEORY, T(17: 1), P*, S, L. *Introduction to the Mathematical Theory of Control Processes, Nonlinear Processes, Volume II.* Richard Bellman. Acad Pr, 1971, xix + 306 pp, $16. This second volume of three planned contains theoretical and computational aspects of continuous and discrete non-linear

control processes and their bases in the calculus of variations and dynamic programming. It is written in the author's usual, enjoyable style with short sections, informative exercises, and annotated bibliography. It purposely stops short of Hamilton-Jacobi theory and actual applications and concludes with a chapter on directions for research. R.W.N.

APPLIED MATHEMATICS, CONTROL THEORY, S, P, L. *Optimal Control of Systems Governed by Partial Differential Equations.* J.L. Lions. Transl: S.K. Mitter. *Die Grundlehren der Mathematischen Wissenschaften, Band 170.* Springer-Verlag, 1971, xi + 396 pp, $22.60. This monograph, translated from the 1968 French edition, discusses the control of systems in which the state is a solution to a partial differential equation. Included, among other considerations, are existence, necessary and sufficient conditions depending upon the classification of the partial differential operator, and regularization, approximation and penalization. R.W.N.

APPROXIMATIONS, P, L. *Best Approximation in Normed Linear Spaces by Elements of Linear Subspaces.* Ivan Singer. Transl: R. Georgescu. *Die Grundlehren der Mathematischen Wissenschaften, Band 171.* Springer Verlag, 1970, 415 pp, $17.30. The methods of functional analysis are applied to the problems of best approximation in normed linear spaces. Major chapters are devoted to approximations by elements from arbitrary linear subspaces, linear subspaces of finite dimension, and closed linear subspaces of finite codimension. Considered are characterizations, existence, and uniqueness of solutions with applications to specific spaces. Appendices are devoted to approximations by elements of convex sets, surfaces, and arbitrary sets. R.B.K.

COMPLEX ANALYSIS, T(16-17: 1). *Techniques of Asymptotic Analysis. Applied Mathematical Sciences, Volume 2.* L. Sirovich. Springer-Verlag, 1971, ix + 306 pp, $6.50 (P). This book is in the new *Applied Mathematical Sciences* series, a collection of paperbacks printed directly from typescript and written for those with interests in applications of mathematics. This volume is as concerned with computational techniques in asymptotic analysis as it is with theory, and it aims to develop ability in formal reasoning. Illustrations, reading list, many exercises. Chapter titles: Asymptotic Sequences and Asymptotic Development of a Function, The Asymptotic Development of a Function Defined by an Integral, Linear Ordinary Differential Equations. D.F.A.

DIFFERENTIAL AND INTEGRAL EQUATIONS, T(18), P, L. *Differential and Integral Inequalities.* Wolfgang Walter. Transl: Lisa Rosenblatt. *Ergebnisse der Mathematik und ihrer Grenzgebiete, Band 55.* Springer-Verlag, 1970, x + 352 pp, $21.40. A translation of the author's 1964 monograph, but with new results and some additional topics, including the line method for parabolic equations. The author uses differential and integral inequalities to attack problems in the theory of differential and integral equations; he examines Volterra integral equations (in one and several variables), ordinary differential equations, and hyperbolic, parabolic, and (in an appendix) elliptic partial differential equations. Many examples and an extensive bibliography. D.F.A.

DIFFERENTIAL EQUATIONS, T*(14: 1). *Ordinary Differential Equations.* Otto Plaat. Holden-Day, 1971, xi + 295 pp, $10.95. The author presents a qualitative approach to differential equations which should enable the student to think about differential equations rather than merely to manipulate them. The student should acquire a firm grasp of the geometric nature of differential equations and the problems associated with them. There is a unique chapter on plane autonomous systems, with numerous examples to illustrate. L.L.K.

DIFFERENTIAL EQUATIONS, STABILITY, T(16), S, P, L. *Matrix Methods in Stability Theory.* S. Barnett and C. Storey. B & N, 1970, xiii + 148 pp, $9.50. Although intended as an introduction for students of mathematics, science, and engineering, this book may be useful also to those looking for a better understanding of stability or applications of matrices. The first half of the book forms the matrix theory background needed for the second half, which consists of a study of solution methods for the Liapunov matrix equation and applications. R.W.N.

FOUNDATIONS, L*, P. *Mathematical Logic and Foundations of Set Theory.* Yehoshua Bar-Hillel. North-Holland, 1970, 145 pp, $9.80. Proceedings of the Congress on "Mathematical Logic and Foundations of Set Theory", Jerusalem, November 1968. M.O. Rabin, Weakly definable relations and special automata. Y.N. Moschovakis, Determinancy and prewellorderings of the continuum. C.E.M. Yates, Initial segments of the degrees of unsolvability. R.B. Jensen and R.M. Solovay, Some applications of almost disjoint sets. H. Gaifman, On local arithmetical functions and their applications for constructing types of Peano's arithmetic. R. Jensen, Definable sets of minimal degree. A. Levy, Definability in axiomatic set theory II. J.G.L.

FOUNDATIONS OF QUANTUM THEORY, P. *Quantum Theory and Beyond: Essays and Discussions Arising From A Colloquium.* Ed: Ted Bastin. Cambridge U Pr, 1971, ix + 345 pp, $16. A collection of twenty-three papers dealing with the mathematical and philosophical problems arising from the circularities and apparent contradictions in Quantum Theory. Some interesting new models are proposed. T.A.V.

FUNCTIONAL ANALYSIS, T(17: 1), S. *Topics in Operator Theory.* Richard Beals. U of Chicago Pr, 1971, x + 130 pp, $2.50 (P). Lecture notes developed in a course at the University of Chicago "The aim...was to proceed from a basic knowledge of bounded linear operators in Hilbert space to some of the deeper and more interesting parts of the theory of linear operators." T.A.V.

FUNCTIONAL ANALYSIS, T(18: 2), P, L. *Harmonic Analysis of Operators on Hilbert Space.* Béla Sz.-Nagy and Ciprian Foias. North-Holland, 1970, xiii + 387 pp, $20.75. "The purpose of the present monograph is to give a detailed exposition of the information about a contraction which can be obtained from its unitary dilation." The authors presume a knowledge of the elements of the theory of Hilbert spaces at the level of the text by F. Riesz and B. Sz.-Nagy, and familiarity with H_p spaces. Contains an extensive bibliography. T.A.V.

GENERAL, T(13: 1). *Mathematics: Art and Science.* S.M. Dowdy. Wiley, 1971, xviii + 282 pp, $8.95. For nonmathematicians. A beautiful attempt to present mathematics as a creative art and still capture its role in science. Nine chapters develop a very readable introduction to many branches of mathematics: Number Theory, Modern Algebra, Geometry, Foundations, etc. Suggested readings follow each chapter. L.L.K.

GENERAL, T(1), L. *Mathematics: A Human Endeavor.* Harold R. Jacobs. Freeman, 1970, xvii + 529 pp, $8.50. The content could hardly be classified as college level material but the approach is very refreshing and deserves special attention. Over two-thirds of the book consists of carefully chosen questions, exercises and experiments, graded for difficulty, which will appeal to "those who think they don't like the subject," and which emphasize inductive thinking and discovery. Topics include number sequences, functions and graphs, large numbers and logarithms, polygons and conics, probability and statistics, and topology. 634 illustrations! L.C.L.

GENERAL, T(13: 1, 2). *The Calculus Book: A First Course With Applications and Theory.* Louis Leithold. Har-Row, 1971, v + 853 pp, $12.95. Another calculus text you might consider for a course given for poorly prepared students. There are more applications to economics than usual, but still it seems that there is no need to hurry your order for an examination copy since the author's 1969 2-volume version of calculus contains almost all the other material and much more. W.C.R.

GENERAL, T(13: 2). *Applied Mathematics: An Introduction. Mathematical Analysis for Management.* Chris A. Theodore. Richard D. Irwin, 1971, xiii + 722 pp, $12.95. For nonmathematicians. This is the second edition of a book first printed in 1965. There are many additions, deletions, and revisions. Part I consists of 7 sections on sets and logic; Part II, fundamentals of algebra and analytic geometry; Part III, functions and their application to business operations; and Part IV, elements of calculus with applications. This last part is 200 pages in length and deals with only those parts of calculus which the author deems essential for business. L.L.K.

GENERAL, T(13), S, L. *Elementary Vectors, Second Edition.* E. OE. Wolstenholme. Pergamon Pr, 1971, vii + 109 pp, $2.35 (P). Too brief for use as a text. (Only 84 small pages plus problems.) Material is well organized. It is designed to cover the requirements of certain syllabi in the British educational system. K.W.

GENERAL, P, L. *Lecture Notes in Mathematics-179: Séminaire Bourbaki, Volume 1968-69, Exposés 347-363.* Springer-Verlag, 1971, iv + 295 pp, $6.40 (P), and *Lecture Notes in Mathematics-180: Séminaire Bourbaki, Volume 1969-70, Exposés 364-381.* iv + 310 pp, $6.10 (P). Two volumes continuing an important series. (Exposés 1-346 reprinted by Benjamin in 15 volumes, covering the years 1948-1968.) Varied topics. L.A.S.

GENERAL APPLICATIONS, B, L. *Explorations in Mathematical Anthropology.* Ed: Paul Kay. MIT Pr, 1971, xviii + 286 pp, $12. A collection of 14 papers on mathematical, statistical, and computer

techniques in anthropology. All but 3 of the papers were presented in 1966 to the AAAS (Section H) meetings in Berkeley. The book's goal is to show anthropologists that mathematics can be used to do significant work. This book might be suggested for a joint seminar, but not a math course. W.C.R.

GENERAL, DICTIONARY. *Vocabularie Mathematic.* C.E. Sjöstedt. Interlingue Uppsala, 1970, 87 pp. The main vocabulary list gives Interlingue words and their English, French, and German synonyms. Three separate briefer lists refer words in these languages to Interlingue. A perennial obstacle to the use of such constructed languages remains unsurmounted: there aren't enough words. The term "fasce" is indifferently offered as meaning pencil, bundle, and sheaf. Predictably, we have campe = field = champe = Feld. Of limited value. L.A.S.

GENERAL, LINGUISTICS, T(16-17: 1), P, L. *Introduction to the Mathematics of Language Study. Number eight in the series Mathematical Linguistics and Automatic Language Processing.* Barron Brainerd. Am Elsevier, 1971, x + 313 pp, $18. For readers knowing some linguistics but no mathematics, the book presents mathematical concepts in the contexts in which they are used in the study of linguistics. The first three chapters discuss set theoretic models for linguistic structures, and the final two generative--production and transformational-grammars. Exercises, chapter bibliographies. (Unreasonably expensive considering the format.) L.A.S.

HISTORY, P, L*. *Opera Mathematica.* Francois Viète. Georg Olms Verlag, 1970, lii + 554 pp, $25.15. Facsimile reprint of the original edition edited by Francisci a Schooten (Leiden 1646), plus a 30 page well-footnoted introduction and a 10 page index by Joseph E. Hoffmann. Although we all "know" that Viète introduced modern algebraic notation, it is worth looking at this primary source if only to see how halting and incomplete his innovations appear to us now. K.O.M.

HISTORY, P, L. *Peter und Philipp Apian, zwei deutsche Mathematiker u. Kartographen.* Dr. Siegmund Günther. Meridian, 1967, 136 pp, $5. A facsimile reprint of the first edition (Prague, 1882) of the only biographical study of the sixteenth century German father (Peter) and son practitioners and innovators in arithmetic, trigonometry, cartography, topography, and astronomy. K.O.M.

HISTORY, APPLIED MATHEMATICS, L. *The Universe of the Mind.* George E. Owen. Johns Hopkins Pr, 1971, xiv + 349 pp, $15, $4.95 (P). An historical treatment of ideas which joined math and physics until the early 19th century and also those ideas that caused the fields to diverge during that century. The orientation of the book is strongly toward physics students. There are no problems. W.C.R.

HISTORY, CALCULUS, S, P, L*. *Reflexions sur La Metaphysique du Calcul Infinitesimal.* Lazare Carnot. Albert Blanchard, 1970, xiv + 153 pp, 12F. Reprint with an eight page biographical note (by Marcel Mayot) of the work of 1797 which C.B. Boyer calls "perhaps the most famous attempt to clear up the difficulties" in the foundations of the calculus *(The History of the Calculus*, p. 257). Indeed this unsuccessful effort went through many French editions, was translated into English, Portugese, German and Italian, and is

still interesting reading for anyone interested in the foundations of analysis. K.O.M.

HISTORY, CAUCHY, P, L. *La Vie Et Les Travaux du Baron Cauchy.* C.-A. Valson. Albert Blanchard, 1970, xxxii + 178 pp, $5.90. A reprint of the only full length biography, first published in 1868 by the editor of Cauchy's collected works and here accompanied by a critical introduction by Rene Taton, the dean of French historians of mathematics, who urges further study of the scientific work of this towering nineteenth century figure. K.O.M.

LINEAR ALGEBRA, T(13-14: 1). *Linear Algebra With Differential Equations.* Sze-Tsen Hu. Markham, 1971, xiii + 374 pp, $8.95. "With differential equations" means in this case an appended chapter, as an application, to a book on linear algebra. It is an elementary introduction to differential equations, not a combined study of the two topics. L.L.K.

MEASURE THEORY, T(17-18: 2), L. *Measure and Integration.* Sterling K. Berberian. Chelsea, 1965, xviii + 312 pp, $7.95. Reprint of 1965 edition. Preliminary edition was *A First Course in Measure and Integration*, 1962. The basic theory of measure and integration over abstract measure spaces is followed by chapters on product measures, finite signed measures, integration over locally compact spaces, and integration over locally compact groups. Careful treatment of null sets is recognized as the key to clarity. Organizations makes reference easy. Complete index of symbols. R.B.K.

MEASURE THEORY, T(17-18: 1), P, L. *Hausdorff Measures.* C.A. Rogers. Cambridge U Pr, 1970, viii + 179 pp, $12.50. Hausdorff measures are defined on a metric space with respect to monotonic right continuous functions on the non-negative reals, and generalize the notion of r-dimensional area in n-space. (In a certain sense, the Cantor ternary set has dimension log 2/log 3.) After an introduction to measure theory, the general theory of Hausdorff measures is developed. Applications are surveyed. Good documentation and bibliography. R.B.K.

POTENTIAL THEORY, T(18: 1), P. *An Introduction to Potential Theory. University Mathematical Monographs.* Nicolaas du Plessis. Hafner, 1970, viii + 177 pp, $10.95. A mathematically sophisticated treatment of potential theory based on the Lebesgue-Stieltjes integral. Preliminaries; superharmonic, subharmonic, and harmonic functions in R^p; the conductor problem and capacity; the Dirichlet Problem (in locally compact Hausdorff spaces). References. No exercises. R.B.K.

Reviewers Whose Initials Appear Above

David F. Appleyard, Carleton; John Dyer-Bennet, Carleton; Lorraine L. Keller, St. Olaf; Roger B. Kirchner, Carleton; Loren C. Larson, St. Olaf; John G. Lewis, St. Olaf; Kenneth O. May, University of Toronto; R.W. Nau, Carleton; William C. Ramaley, Carleton; Linda A. Seebach, St. Olaf; T.A. Vessey, St. Olaf; Kenneth Wegner, Carleton.

NEWS AND NOTICES

EDITED BY RAOUL HAILPERN, SUNY at Buffalo

Readers are invited to contribute to the general interest of this Department by sending news items to Mathematical Association of America, 1225 Connecticut Avenue, N.W., Washington, D. C. 20036. Items must be submitted at least two months before publication can take place.

PERSONAL ITEMS

Dr. E. R. Berlekamp, Bell Laboratories, has won honorable mention from Eta Kappa Nu Engineering Honor Society as one of the nation's three outstanding young electrical engineers.

Dr. Herbert Busemann, Distinguished Emeritus Professor at the University of Southern California, received an honorary degree of Doctor of Laws at USC's sixth Mid-year Commencement Exercises.

Professor Morris Kline, New York University, received one of the three NYU's Great Teacher Awards for 1971.

Professor R. S. Varga, Kent State University, has been commissioned an honorary Kentucky Colonel.

Boston College: Associate Professor G. G. Bilodeau has been promoted to Professor; Assistant Professor R. L. Faber has been promoted to Associate Professor.

Associate Dean G. F. Clanton, Vanderbilt University, has been appointed Associate Provost and Dean for Academic Planning.

Dr. A. N. Feldzamen, Encyclopaedia Britannica Educational Corporation, has been appointed Vice President and Editorial Director, Films and Publications.

Associate Professor Anthony Mardellis, California State College, Long Beach, has been promoted to Professor.

Professor E. E. Moise, Harvard University, has been appointed to a Distinguished Professorship at Queens College (CUNY).

Mr. Frank Montgomery, Brescia College, has been appointed Assistant Professor and Chairman of the Mathematics Department.

Professor S. A. Naimpally, Indian Institute of Technology, Kanpur, has been appointed to a Visiting Professorship at Lakehead University.

Professor R. R. Stoll, Oberlin College, has been appointed Professor and Chairman of the Department of Mathematics at Cleveland State University.

Professor Emeritus Helen Barton, University of North Carolina, Greensboro, died on March 19, 1971 at the age of 79. She was a Charter Member of the Association.

Dean Emeritus Samuel Beatty, University of Toronto, died on July 30, 1970 at the age of 88. He was a member of the Association for fifty-three years.

Professor Emeritus A. A. Bennett, Brown University, died on February 17, 1971 at the age of 82. He was a Charter Member of the Association and a former Editor of the MONTHLY.

Professor Louis Brand, University of Houston, also Professor Emeritus of the University of Cincinnati, died on January 27, 1971 at the age of 85. He was a Charter Member of the Association.

Dr. H. K. Brown, Somerville, Massachusetts, died on August 9, 1970 at the age of 55. He was a member of the Association for twenty-nine years.

Dr. Joseph Golob, Dayton, Ohio, died on September 25, 1970 at the age of 65. He was a member of the Association for eighteen years.

Professor P. L. Meyer, Washington State University, died on November 27, 1970 at the age of 44. He was a member of the Association for seven years.

Professor Emeritus W. E. Milne, Oregon State University, died on January 19, 1971 at the age of 81. He was a Charter Member of the Association.

VISITING LECTURER PROGRAM IN STATISTICS

A Visiting Lecturer Program in Statistics has been organized for the ninth successive year. The program is sponsored jointly by the principal statistical organizations in the United States, the American Statistical Association, the Biometric Society, and the Institute of Mathematical Statistics. The National Science Foundation provides partial financial support. Leading teachers and research workers in statistics—from universities, industry, and government—have agreed to participate as lecturers. Lecture topics include subjects in experimental and theoretical statistics, as well as in such related areas as probability theory, information theory, and stochastic models in the physical, biological, and social sciences.

The purpose of the program is to provide information to students and college faculty about the nature and scope of modern statistics, and to provide advice about careers, graduate study, and college curricula in statistics. Inquiries should be addressed to: Visiting Lecturer Program in Statistics, Department of Statistics, Southern Methodist University, Dallas, Texas 75222.

The list of participating lecturers includes: Z. W. Birnbaum (University of Washington), M. W. Carter (Brigham Young University), Arthur Cohen (Rutgers University), T. M. Cover (Massachusetts Institute of Technology), E. L. Crow (U. S. Department of Commerce Research Laboratories, Boulder, Colorado), H. A. David (University of North Carolina), H. T. David (Iowa State University), W. A. Ericson (University of Michigan), D. A. Gardiner (Oak Ridge National Laboratory), J. L. Gastwirth (Office of Statistical Policy, Washington, D. C.), J. P. Gilbert (Harvard University), L. J. Gleser (Johns Hopkins University), R. Gnanadesikan (Bell Telephone Laboratories, Murray Hill, New Jersey), W. C. Guenther (University of Wyoming), H. L. Harter (ARL, Wright-Patterson AFB, Ohio), L. H. Herbach (New York University), C. R. Hicks (Purdue University), W. M. Hirsch (New York University), R. R. Hocking (University of Houston), Myles Hollander (Florida State University), P. W. M. John (University of Texas at Austin), Leo Katz (Michigan State University), S. Kotz (Temple University), C. Y. Kramer (Virginia Polytechnic Institute), I. H. LaValle (Tulane University), Eugene Lukacs (The Catholic University of America), B. J. Mandel (Washington, D. C.), C. E. Marshall (Stillwater, Oklahoma), John Neter (University of Minnesota), P. L. Odell (Texas Tech University), S. J. Press (University of Chicago), Tim Robertson (University of Iowa), R. N. Schmidt (SUNY at Buffalo), Paul Switzer (Stanford University), D. L. Sylwester (University of Washington), B. J. Tepping (Washington, D. C.), W. A. Thompson, Jr. (University of Missouri), B. J. Trawinski (University of Alabama at Birmingham), B. E. Trumbo (California State College at Hayward), Grace Wahba (University of Wisconsin), K. T. Wallenius (Clemson University), J. W. Wilkinson (Rensselaer Polytechnic Institute), James Yackel (Purdue University).

The organizing committee consists of: H. T. David, S. W. Greenhouse, S. S. Gupta, W. J. Hall, L. Katz, L. H. Koopmans, I. Olkin, D. B. Owen (Chairman), G. J. Resnikoff.

THE MATHEMATICAL ASSOCIATION OF THE UNITED KINGDOM

The Mathematical Association of the United Kingdom (M.A.), founded in 1871, lists amongst its objectives the improvement of the teaching of mathematics and its applications, and the provision of means of communication between students and teachers of mathematics.

The Association promotes such communication through its meetings—both at national and branch level—and, in particular, through its publications: *The Mathematical Gazette*, *Mathematics in School*, its *Newsletter*, and its various Reports.

The Mathematical Gazette was first published in 1894 and now appears 4 times a year, in February, May, October, and December. The Gazette contains articles, classroom notes, and correspondence, dealing with the teaching of mathematics as well as with mathematical topics of general interest, and also features a large number of reviews of text books and of publications of a more advanced type from many British and other publishers. In the past the *Gazette* has attempted to cover the whole range of school mathematics. In the future, however, it is likely to concentrate more on the teaching of mathematics to students of 16 years and over, for a new periodical, *Mathematics in School*, will have as its primary concern the teaching of mathematics to pupils in the 9–16 age range. This new periodical will be published six times a year (in the odd-numbered months) and will be intended for the general teacher of mathematics rather than the specialist.

The *Newsletter* appears some two or three times a year and contains notices of events, etc., of general, yet ephemeral, interest.

All members of the Association receive free copies of the Reports prepared by the Association's Teaching Committee. There have been some seventy such reports since the first—on the teaching of geometry—was published in 1902, and they represent a major part of the Association's activities. Amongst the eleven reports published since 1968 have been *Mathematical Laboratories in Schools*, *Mathematics Projects in British Secondary Schools*, *Computer Studies in Schools*, *Primary Mathematics: A further report*, *The Same but Different* (a survey of the notion of equivalence in the context of school mathematics), and *Computers and the Teaching of Numerical Mathematics in the Upper Secondary School.*

Members of the Association may elect to receive either or both of the periodicals and all are sent copies of reports and newsletters. Under the reciprocity agreement subscribed to by the Mathematical Association of the United Kingdom and the Mathematical Association of America, members of the latter association may become members of the M.A. at the following (reduced) rates:

$12 per year—*Mathematical Gazette and Mathematics in School*
$ 8.75 per year—*Mathematical Gazette* only.

Membership is for the calendar year and dues are payable in advance on January 1st. (It is planned to publish the first issue of *Mathematics in School* in Autumn 1971, and a copy of this first issue will be sent to all those who subscribe to receive the periodical in the year 1972.) Dues should be paid to the Washington Office of the MAA from which office an application form for membership can be obtained.

The special rates given above will also apply to those members of the MAA who are currently members of the M.A. In order to obtain these rates, such members should pay their dues direct to Washington and should also inform the Executive Secretary of the Mathematical Association (150 Friar Street, Reading, RG1 1HE, England) of the change in their mode of payment.

13TH INTERNATIONAL CONGRESS OF THEORETICAL AND APPLIED MECHANICS

The 13th Congress will be held at Moscow, U.S.S.R., from Monday 21 August to Saturday 26 August 1972. The meetings and lectures will take place at Moscow State University. Accommodation for participants will be available in hotels in the city, and a limited number of participants can be accommodated in the hostels of the University.

The 13th Congress will encompass the entire field of the science of analytical, solid and fluid mechanics, including applications. Computational methods as such will not be included. There will be a number of general and sectional lectures given by speakers on the invitation of the International Program Committee, and between 200 and 250 contributed papers will be presented.

Approximately 50 of the papers contributed by residents of the United States will be accepted. An initial selection from U.S. contributions will be made by the U.S. National Committee on Theoretical and Applied Mechanics of the National Academy of Sciences—National Research Council on the basis of abstracts to be refereed by a broadly representative national selection committee. The final selection of U.S. papers will subsequently be made by the International Program Committee. *Each contributor of a paper from the U.S.A. should send 5 copies of a summary of 500 words to reach Professor G. F. Carrier, Pierce Hall, Harvard University, Cambridge, Mass. 02138, before March 15, 1972.*

People who are potentially interested in taking part in the Congress may receive further details by expressing their interest to the General Secretary of the Organizing Committee at the following address: Professor G. K. Mikhailov, Leningradskii Prospekt 7, Moscow A-40, U.S.S.R.

ACM CONFERENCE ON PROVING ASSERTIONS ABOUT PROGRAMS

The ACM Special Interest Groups on Programming Languages (SIGPLAN) and Automata and Computability Theory (SIGACT) will jointly sponsor a conference on Proving Assertions about Programs on January 6–7, 1972. The program will include both contributed and invited papers. The conference will be held at New Mexico State University which is located at Las Cruces, New Mexico, less than an hour's drive from El Paso, Texas, and Juarez, Mexico. Appropriate topics include, but are not limited to, design of languages to facilitate proofs, relationship of formal language definition to proofs of assertions, equivalence to problems of logic, implications of undecidability results, proof methods based on induction.

Enquiries concerning contributed papers should be addressed to R. H. Stark, Computer Science Department, New Mexico State University, Las Cruces, New Mexico 88001. Registration inquiries should be addressed to the Conference Chairman, J. Mack Adams, Computer Science Department, New Mexico State University, Las Cruces, New Mexico 88001.

NCTM ANNUAL MEETING, CHICAGO, ILLINOIS, APRIL 16–20, 1972

A call for research papers in mathematics education to be read at this meeting has been made. Research on the teaching of mathematics at the college and university level, as well as precollege levels, is especially welcomed. Proposals for research papers must be submitted by December 31, 1971. For details, write to Professor Jon L. Higgins, ERIC Information and Analysis Center for Science and Mathematics Education, 1460 West Lane Avenue, Columbus, Ohio 43210.

MATHEMATICAL ASSOCIATION OF AMERICA

Official Reports and Communications

NOVEMBER MEETING OF THE MARYLAND-DISTRICT OF COLUMBIA-VIRGINIA SECTION

The annual Fall meeting of the Maryland-District of Columbia-Virginia Section of the MAA was held at Northern Virginia Community College in Annandale, Virginia on November 21, 1970. Mr. William Norris, Chairman of the Section, presided at the meeting. A total of 123 persons registered at the meeting, 95 of whom were members.

The following papers were presented:

Space filling pentahedra, by Michael Goldberg, Washington, D. C.

Peano-like postulates for natural numbers, by R. F. McCoart, Loyola College, Baltimore, Maryland.

Polynomial identities in non-commutative algebras, by Jerry Karaganis, American University, Washington, D. C.

Behavioral objectives and the teaching of mathematics, by J. M. Smith, George Mason College, Fairfax, Virginia.

Self-paced learning of calculus, by Arnold Stokes, Georgetown University, Washington, D. C.

Generalized functions in elementary calculus?, by R. L. Eiseman, USAF and University of Maryland, College Park, Maryland.

Dr. Ruth Bari, George Washington University, presented an invited paper entitled, "A New Look at the Four-Color Conjecture."

MARY CABELL, *Secretary*

NOVEMBER MEETING OF THE PHILADELPHIA SECTION

The forty-fifth annual meeting of the Philadelphia Section of the MAA was held at West Chester State College, West Chester, Pennsylvania on November 21, 1970. The Section Chairman, Professor Willard Baxter of the University of Delaware, presided at the meeting, which was attended by 225 persons, including 172 members of the Association.

The following business was transacted:

1. The following officers were elected: Chairman, Hugh Albright, LaSalle College; Vice-Chairman, W. J. Pervin, Drexel University; Secretary-Treasurer, A. E. Filano, West Chester State College; Executive Committee Members-at-Large, James Brooks, Villanova University; Joseph Mamelak, Community College of Philadelphia.

2. The two top performers from the Section in the 1969 Putnam Competition were recognized and awarded an annual membership in the MAA. They are: A. E. Barnes, Drexel University; B. J. Kuipers, Swarthmore College.

Honorable mention citations were also presented to: T. C. Folsom, Villanova University; Kathy L. Kasley, Ursinus College; Kay Pechenick, University of Pennsylvania.

The following papers were presented:

1. *Some unsolved problems from intuitive geometry*, by V. L. Klee, University of Washington.
2. *What is an FK space?* by Albert Wilansky, Lehigh University.
3. *Computers and the first two years of college mathematics.* Panelists: P. J. Davis, Brown University; Gerald Porter, University of Pennsylvania; L. C. Leinbach, Gettysburg College.
4. *Accreditation and certification in mathematics.* Panelists: Barnard Bissinger, Pennsylvania State University; W. E. Baxter, University of Delaware; W. R. Jones, Lafayette College.

A. E. FILANO, *Secretary-Treasurer*

NOVEMBER MEETING OF THE UPPER NEW YORK STATE SECTION

The fall meeting of the Upper New York State Section of the MAA was held at the State University College at Oswego, New York, on November 7, 1970. There were 92 members and 23 guests in attendance. Professor R. Sloan, Chairman of the Section, presided at the sessions.

A panel discussion by S. O. Hockett, Ithaca College, D. E. Kibbey, Syracuse University, and F. R. Olson, State University College at Fredonia, considered the topic "Accreditation and Certification in Mathematics."

In addition, the following papers were presented:

1. *Embedding a partially ordered ring in a division ring*, by W. H. Reynolds, State University at Cortland.
2. *Simple proofs of elementary probability theorems by use of indicator functions*, by P. Tan, Carleton University.
3. *Generalized generalized inverses*, by G. Rabson, Clarkson College of Technology.

The meeting concluded with an invited address, "Topological methods in analysis," by G. S. Young, University of Rochester.

P. SCHAEFER, *Secretary-Treasurer*

FEBRUARY MEETING OF THE NORTHERN CALIFORNIA SECTION

The annual meeting of the Northern California Section of the MAA was held at the University of San Francisco on February 6, 1971; one hundred and seventy people were in attendance.

Professor John Fischer, S.J., Chairman of the Mathematics Department, University of San Francisco, opened the meeting with greetings of welcome, and Mary V. Sunseri, program chairman, presided at the morning session, during which the following papers were presented:

1. *Matijasevic's answer to Hilbert's 10th problem*, by Julia B. Robinson, University of California, Berkeley.
2. *Some mathematical problems associated with the study of elections*, by James Dolby, San Jose State College.

At the business meeting, the following officers were elected: Chairman, G. L. Alexanderson, University of Santa Clara; Vice-Chairman, Leonard Tornheim, Chevron Research; Secretary-Treasurer, N. H. Fisher, San Francisco State College; Program Chairman, T. H. Southard, California State College, Hayward.

Professor T. H. Southard, chairman of the section, presided at the afternoon session which included the following program:

3. *What is computer science anyway?*, by G. E. Forsythe, Stanford University.
4. *The junior college five years later*, by R. L. Norman, Dartmouth College.
5. *Panel discussion: Accreditation and certification*. Members of panel: L. H. Lange, San Jose State College; D. W. Bushaw, Washington State University; Joseph Hashisaki, Western Washington State College; G. B. Pedrick, California State College, Hayward.

N. H. FISHER, *Secretary-Treasurer*

MARCH MEETING OF THE FLORIDA SECTION

The fourth annual spring meeting of the Florida Section of the MAA was held on March 19 and 20, 1971, at Florida Southern College in Lakeland, Florida.

Eleven invited addresses were presented as follows: "A Survey of Elementary Homological Algebra," by Professor Alex Rosenberg, Cornell University; "Accreditation and Certification in Mathematics," by Professor Daniel Finkbeiner, Kenyon College; "Composite Functional Equations," by Professor A. D. Wallace, University of Florida; "On Generalizations of Limit of a Sequence," by Professor D. L. Sherry, University of West Florida; "Undergraduate Program in Probability and Statistics," by Professor Richard Cornell, Florida State University; "The Use of Games in Teaching Mathematical Thought Processes," by Professor H. W. Thwing, Stetson University; "Independent Sets and Complete Subgraphs," by Professor A. R. Bednarek, University of Florida; "What is an Algebraic Group?," by Professor David Hertzig, University of Miami; "Absolutely Pure Modules—The Development of an Idea," by Professor B. H. Maddix, Florida Presbyterian University; "On Certain Improperly Posed Problems for Hyperbolic Equa-

tions," by Professor E. C. Young, Florida State University; "Graph Sequences," by Professor A. W. Goodman, University of South Florida.

Professors Finkbeiner and Rosenberg held a panel discussion on Accreditation and Certification. Professors Jones and Whitman presented a report on the Laramie Conference on Math. for Developing Colleges.

In conjunction with the meeting there was a State Articulation Conference which began with a report on progress in articulation and an open discussion on current questions.

Dr. Bill Gager, Jr., discussed how the State Department of Education looks at articulation in Florida. Professor Eugene Nichols, Florida State University, conducted a panel discussion on the Mathematics Content for the Elementary Education Major. Professor Alex Rosenberg conducted a panel discussion on CUPM Recommendations. Professor Daniel Finkbeiner discussed the various roles of statistics and linear algebra in the two-year college. Professor Howard Taylor of West Georgia College presented: "A Report on an Experiment with General Education Mathematics at West Georgia College." Professor Lewis Edwards of Valencia Junior College presented the COSIPS Report.

The following papers were presented:

1. *Dimension theory in power series rings*, by D. E. Fields, Stetson University.
2. *A metric characterization of zero-dimensional spaces*, by Ludvik Janos, University of Florida.
3. *How to avoid topology in Hilbert spaces*, by Robert Piziak, University of Florida.
4. *Non-Hausdorff sequential convergences*, by J. M. Anthony, Florida Technological University.
5. *A study of attitude changes in college freshmen mathematics students*, by Marilyn L. Repsher, Jacksonville University.
6. *An experiment in teaching general education mathematics*, by R. D. Hackworth, St. Petersburg Junior College, Clearwater Campus.
7. *Note on alternating series and introduction to the addition formulas*, by J. F. Golightly, Jacksonville University.
8. *Some new results of research comparing programmed and lecture-text instruction*, by M. E. Nott, Jr., St. Petersburg Junior College.
9. *A property of 30 and its generalization*, by Hermann Simon, University of Miami.
10. *On multiplicative completion of certain basic sequences in L^p spaces*, by Ben-Ami Braun, University of South Florida.
11. *Is convergence in probabilistic metric spaces topological?*, by Howard Sherwood, Florida Technological University.
12. *The generalized complementarity problem*, by A. L. Price, University of South Florida.
13. *L^2 convergence of interpolation polynomials on the unit circle*, by P. J. O'Hara, Florida Technological University.
14. *A geometric explanation of nutation*, by A. D. Snider, University of South Florida.
15. *$N^{p/q}$, an infinite product*, by J. F. Golightly, Jacksonville University.
16. *Vector methods for high school students*, by Charles McCracken, Admiral Faragut Academy.
17. *Do students learn from and like an audio tape course in freshman mathematics?*, by P. M. Wilson, Florida A&M University.
18. *Oscillation of solutions of third order nonlinear differential equations*, by W. M. Wanamaker, University of South Florida.

Professor Gene Medlin presided at the business meeting following a luncheon at noon Saturday. The following officers were elected for 1971–72: Chairman, Professor Charles McArthur, Florida State University; Chairman-elect, Professor Paul McDougle, University of Miami; Vice-Chairman, Professor George Cash, Manatee Junior College; Secretary-Treasurer, Professor Frank Cleaver, University of South Florida.

F. L. CLEAVER, *Secretary*

MARCH MEETING OF THE OKLAHOMA-ARKANSAS SECTION

The 33rd annual meeting of the Oklahoma-Arkansas Section of the MAA was held at the University of Tulsa, Tulsa, Oklahoma on March 12–13, 1971, with 114 members and 201 total persons in attendance. The Chairman, Professor Raymond McKellips of Southwestern State College, presided over the meeting.

A dinner was held on the evening of March 12 followed by a film entitled "John von Neumann." Four calculus films entitled "Maximize," "Infinite Acres," "The Definite Integral," and "Newton's Method" were shown Saturday morning. Dr. Ivo Babuska gave the invited address entitled "Numerical Methods of Solving Elliptic Partial Differential Equations." A panel discussion on Accreditation and Certification in Mathematics was given by Professors L. Mason, chairman, T. McKellips, R. Reynolds, J. Jewett, W. Orton, and G. Levy.

The following officers were elected: Chairman, Professor Tom Cairns, University of Tulsa; First Vice-Chairman, Professor David Moon, State College of Arkansas; Second Vice-Chairman, Professor Bill Spicer, Oklahoma Military Academy; and Secretary-Treasurer, E. K. McLachlan, Oklahoma State University. Approximately 4223 students participated in the high school mathematics contests, Professor Lysle Mason reported. Professor Raymond McKellips was elected to be the new administrator of the high school mathematics contests.

The following papers were presented:

1. *Szego polynomials in several complex variables*, by J. H. Justice, University of Tulsa.
2. *On Blaschke products*, by J. R. Choike, Oklahoma State University.
3. *Perfect operators*, by J. A. Cisneros, Oral Roberts University.
4. *On some characterization of functions by means of a generalized convolution*, by I. I. Kotlarski, Oklahoma State University.
5. *Some remarks concerning the extension problem for groups in order 2^n*, by J. R. Talburt, University of Arkansas.
6. *The group of lattice automorphisms of the lattice of convergence structures*, by Carroll Riecke, Cameron State College.
7. *The structure of the ring of p-adic integers using methods of abelian groups*, by R. C. Knapp, Oklahoma State University.
8. *Inequalities*, by Douglas Foster, Oral Roberts University.
9. *Computer game strategy*, by R. D. Morton, Oklahoma Christian College.
10. *Continuous and discrete triangular distributions and Monte-Carlo simulation*, by Terence Aitken, Northwestern State College.
11. *Convergent finite difference schemes for nonlinear parabolic equations*, by A. C. Reynolds, Jr., University of Tulsa.
12. *The basilar membrane as a uniformly loaded plate clamped on two spiral boundaries in a plane or on two helicalspiral boundaries*, by H. M. Lieberstein, Wichita State University.
13. *On certain classes of planar dynamical systems*, by R. A. Knight, Oklahoma State University.
14. *On the existence of Banach spaces whose duals are abstract L spaces*, by J. B. Bednar, The University of Tulsa, and H. E. Lacey, The University of Texas at Austin.
15. *The convex cone of $\mathcal{P}_n$ functions*, by M. W. Jeter, Oklahoma State University.
16. *Some remarks about semi-inversion of mappings on linear spaces*, by C. G. Moment, Central State College.
17. *Semi-uniquely colorable graphs*, by D. L. Greenwell, Arkansas State University.
18. *Random variables uniformly distributed on compact abelian groups*, by Peter Flusser, Oklahoma State University.
19. *Strong compactness in limit spaces*, by R. J. Gazik, Arkansas State University.

20. *Minimal l_∞ solution of underdetermined linear systems*, by L. F. Kemp, Jr., Amoco Production Research.

21. *Numerical handling of the electrical activity of cells*, by H. M. Lieberstein, Wichita State University.

22. *Frechet's extreme distributions and the transportation problem*, by L. F. Kemp, Jr., Amoco Production Research.

23. *How many permutations of n symbols are the product of r disjoint cycles?* by Roy Fuller, University of Arkansas.

24. *Partition numbers, permutation numbers and binomial coefficients*, by Tetsundo Sekiguchi, University of Arkansas.

25. *Concerning a problem of Yano and Kobayashi*, by R. H. Bowman, Arkansas State University.

26. *Semigroups in which the product of n elements is expressed as that of fewer elements*, by Naoki Kimura, University of Arkansas.

27. *New radicals for associative rings*, by R. L. Tangeman, Arkansas State University.

28. *The CUPM recommendation on basic mathematics for colleges*, by Terral McKellips, Cameron State College.

29. *Procedural aspects of initiating a mathematics course utilizing programmed texts*, by Jerry Smith, Oklahoma Military Academy.

30. *Using the computer as an instructional tool for elementary calculus*, by G. L. Thesing, Oklahoma State University.

E. K. McLachlan, *Secretary-Treasurer*

MARCH MEETING OF THE SOUTHEASTERN SECTION

The University of Alabama was host to the 50th annual meeting of the Southeastern Section of the MAA on March 26–27, 1971. Professors H. E. Taylor, Chairman of the Section, R. L. Plunkett, and J. R. Wesson presided at the general sessions, and Professors Ben Fitzpatrick, Jr., H. C. Miller, Jr., J. D. Neff, Joe Neggers, and H. V. Park presided at the sessions for contributed papers. Invited addresses were given by Professor W. E. Jenner (Vice-Chairman of the Section) of the University of North Carolina and Professor F. T. Birtel of Tulane University. Professor Jenner's topic was "The present situation in non-associative algebras," and Professor Birtel spoke on "Functionals and algebraic approximation." Under the auspices of CUPM, Professor R. P. Boas of Northwestern University and Professor Donald Bushaw of Washington State University addressed the Section on Saturday morning. Professor Boas reported on the current activities of CUPM and Professor Bushaw discussed accreditation and certification in mathematics. With the cooperation of Modern Learning Aids, two of the MAA films were shown. A total of 227 registered at the meeting.

Officers elected for 1971–72 were: Chairman, R. L. Plunkett, University of Alabama; Chairman-Elect, Trevor Evans, Emory University; Vice-Chairman, H. T. LaBorde, Macon Junior College; and Section Lecturer, C. H. Edwards, Jr., University of Georgia. Professor J. D. Mancill of the University of Alabama was honored by a resolution expressing appreciation for his many years of service to mathematics in our region. Mr. W. H. Beckmann, Jr., of Davidson College received the $25 which is presented to the student in a Southeastern Section school who scores highest in the Putnam Competition.

The following papers were presented:

1. *Non-standard models of the additive group of the integers*, by A. B. Cantor, University of South Carolina.

2. *Multiplicity type and congruence relations in universal algebra*, by T. P. Whaley, Southwestern at Memphis.

3. *M-matrices with respect to a cone*, by L. J. Watford, Jr., Troy State University.
4. *Cosmall modules*, by W. W. Leonard, Georgia State University.
5. *Can we teach students how to study mathematics?*, by J. A. Gore, Valdosta State College.
6. *The effect of prior instruction in logic on understanding limits*, by W. T. Macey, Pfeiffer College.
7. *Simplifying integration by parts*, by R. W. Gibson, Auburn University.
8. *Notation and terminology for the antiderivative*, by J. F. Schell and J. P. Thomas, Western Carolina University.
9. *Evaluation of modern and traditional approaches in teaching freshman mathematics*, Jean B. Mobley, Pfeiffer College.
10. *Points of equicontinuity*, by D. F. Spillman, Guilford College.
11. *Monthly Problem E2271*, by Senior Mathematics Majors of Bennett College, presented by Sharon B. Hudson.
12. *Monthly Problem E2227 and a generalization*, by Joe Albree, University of Tennessee.
13. *Some results in a linearly ordered space with the countable chain condition*, by J. E. Thomas, Guilford College.
14. *Monthly Problem E2278*, by Ruby D. Williams and Shirley J. Sellers, Bennett College.
15. *An expansion theorem for nonanalytic functions in several complex variables*, by M. O. Gonzalez, University of Alabama.
16. *An inversion formula for the Lambert transform*, by E. L. Miller, University of Alabama.
17. *Necessary and sufficient conditions that a function of conjugate complex variables be an exact differential*, by Lola Kiser, Birmingham-Southern College and University of Alabama.
18. *Transform methods for partial differential equations*, by Henry Copeland, University of Alabama.
19. *An algebraic construction for Room squares*, by C. C. Lindner, Auburn University.
20. *A characterization of the group of all order-preserving permutations of a chain*, by E. B. Scrimger, Jr., Southwestern at Memphis.
21. *Two characterizations of semi-hereditary semi-prime Goldie rings*, by John Kinloch, East Tennessee State University.
22. *A counting function of integral n-tuples*, by H. S. Hahn, West Georgia College.
23. *The semigroup of doubly stochastic matrices*, by J. R. Wall, University of Tennessee.
24. *Almost continuous functions dense in the unit square*, by K. R. Kellum, University of Alabama.
25. *Countable regular Hausdorff spaces which are dense-in-themselves*, by J. A. Bond, Jr., Western Carolina University.
26. *Characterizing normal immersions of S^1 into R^2*, by M. L. Marx, Vanderbilt University.
27. *Some remarks on pseudo-expansive mappings*, by L. H. Minor, Western Carolina University.
28. *Example of a piercing set which is not a taming set*, by Harvey Rosen, University of Alabama.

B. F. BRYANT, *Secretary-Treasurer*

MARCH MEETING OF THE SOUTHERN CALIFORNIA SECTION

The 51st regular meeting of the Southern California Section of the MAA was held on March 13, 1971, at San Fernando Valley State College. The registered attendance was 154, including 122 members of the Association. Professor J. A. Ferling, Chairman of the Section, presided.

At the business meeting the elections of Professor Basil Gordon, University of California at Los Angeles as Chairman, and Professor Elmer Tolsted, Pomona College, as Vice-Chairman, for 1971–72 were announced. Professor Robert Herrera, Governor of the Section, reported on the January meeting of the Board of Governors. The following amendment to the Section Bylaws was approved:

"In the event of discontinuation of the Section, the funds in the treasury will revert to the parent organization."

The luncheon speaker was Professor Robion Kirby, U.C.L.A., Veblen Prize Winner.

The following program was presented:

Mathematics and western thought; or What is a group really? by G. C. Henry, Jr., Claremont Men's College.

Mathematical modeling of water pollution in estuaries, by Gene Gritton, Rand Corporation.

Algebraic symmetry in analytic geometry, by R. N. Redheffer, U.C.L.A.

The use of computers in teaching mathematics, by J. W. Bergquist, IBM Corporation.

A constructive Riemann mapping theorem, by E. A. Bishop, U.C.S.D.

Some elementary problems in elementary number theory, by Helmut Hasse, University of Hamburg, currently at San Diego State College.

Panel Discussion: Accreditation and certification. Moderator: J. A. Ferling, Claremont Men's College. Panel Members: L. C. Lay, Cal. State Fullerton, L. J. Paige, U.C.L.A., P. A. White, U.S.C., William Wooton, L. A. Pierce College.

T. N. ROBERTSON, *Secretary-Treasurer*

APRIL MEETING OF THE KENTUCKY SECTION

The fifty-fourth annual meeting of the Kentucky Section of the MAA was held at Western Kentucky University, Bowling Green, on April 2–3, 1971, Chairman R. C. Bueker, Western Kentucky University, presiding. One hundred and nine persons registered for the meeting, including 71 members of the Association.

At the business meeting, Dr. L. L. Scott, University of Louisville, Chairman of the By-Laws Committee, presented a proposed set of By-Laws for the Section and moved the adoption of the set. The motion was seconded and passed. Other members of the Committee were: Dr. B. R. Nail, Morehead State University, and Dr. W. J. Robinson, Centre College of Kentucky, Danville.

After adoption of the By-Laws and in conformity with the new structure, the Nominating Committee, chaired by Dr. Bennie Lane, Eastern Kentucky University, Richmond, presented the following slate of officers: Chairman, Dr. R. A. Dobyns, Georgetown; Chairman-Elect, Dr. J. E. Mack, University of Kentucky, Lexington; Vice-Chairman, Dr. P. E. Bland, Eastern Kentucky University; Secretary-Treasurer, Dr. Walter Feibes, Western Kentucky University. It was moved, seconded and passed that the slate presented be accepted by acclamation.

At 7:00 P.M. on Friday, the mathematics department of Western Kentucky University presented Professor P. R. Halmos, Indiana University, in a Colloquium Address to which members of the Section were invited. Professor Halmos spoke on: *Progress Report on Invariant Subspace Lattices.* He was introduced by Dr. R. R. Crawford, Western Kentucky University.

After the Colloquium Address, the fifty-fourth Annual Meeting began with a panel discussion of "Accreditation /certification of institutions preparing teachers/teachers." Dr. H. G. Robertson, Murray State University, and Dr. J. H. Wells, University of Kentucky, Governor of the Section, comprised the panel.

Saturday morning there were two parallel program sessions, one for contributed mathematics papers and one for contributed mathematics education papers.

Dr. Crawford presided over the session for contributed mathematics papers during which the following papers were presented:

1. *Extensions of mappings in finite Abelian Groups*, by K. D. Wallace, Western Kentucky University.
2. *Selected properties of the Lindelöf functions*, by Amy C. King, Eastern Kentucky University.
3. *Change for a dollar bill*, by Walter Feibes, Western Kentucky University.

4. *Algebraic characterization of some classical combinatorial problems*, by E. T. Ordman, University of Kentucky.

5. *Finite p-Abelian p-Groups*, by R. M. Davitt, University of Louisville.

Mr. W. C. Jones, Western Kentucky University, presided over the mathematics education session during which the following papers were presented:

1. *Randolph diagrams*, by J. B. Barkesdale, Western Kentucky University.

2. *Conceptual mathematical methodology for prospective elementary school teachers*, by A. R. Brousseau, Centre College of Kentucky.

3. *Under-achiever program for 7th and 8th grades*, by a panel comprised of: Ernie Spalding, Bardstown High School, Bardstown; Nancy Henson, University Breckinridge School, Morehead; Johnnie Fryman, Morehead State University.

4. *Computers in secondary mathematics teaching*, by R. H. Geeslin, University of Louisville.

5. *Computers in calculus—a report*, by J. B. Fugate, University of Kentucky, and D. S. Tucker, Morehead State University.

After the two parallel sessions, the two groups met together in the business session which was followed by lunch in the University Center.

Chairman Bueker presided over the afternoon session. Dr. Wells introduced the invited lecturer, Professor Halmos. The title of the lecture was: *Commutators Revisited.*

A. S. HOWARD, *Secretary-Treasurer*

APRIL MEETING OF THE SOUTHWESTERN SECTION

The annual meeting of the Southwestern Section of the MAA was held at Arizona State University, Tempe, Arizona on April 2–3, 1971. Seventy-four persons attended. Professor L. T. Smith, chairman of the section, presided.

The invited guest speaker was Professor Albert Tucker, Princeton University, whose address was titled "Comments on Curricular Reforms." A special panel discussion was held on the problem of "Accreditation and Certification in Mathematics." Panel members were Professors Albert Tucker, Princeton University, Evar Nering, Arizona State University, Don Meyers, University of Arizona, and Robert Wisner, New Mexico State University. Section chairmen were Professors E. E. Grace and P. A. Leonard of Arizona State University, W. W. Mitchell, Phoenix College, and G. S. Rogers, New Mexico State University.

During the business meeting, the following officers were elected: Chairman, Professor S. T. Kao, University of New Mexico; Vice-Chairman, Professor J. F. Foster, University of Arizona; Secretary-Treasurer, Professor R. W. Ball, New Mexico Institute of Mining and Technology.

The following contributed papers were presented:

1. *Loci suggested by a sliding rod*, by Harvey Butchart, Northern Arizona University.

2. *Partial ordering and reflexivity in duality*, by Evar Nering, Arizona State University.

3. *A determination of the finite groups having eight conjugate classes*, by Edwin Annavedder, Arizona State University.

4. *Computer extended calculus and its evolution*, by Gary Bitter, Arizona State University.

5. *The role of dualities in certain classical algebras*, by P. J. Horn, Northern Arizona University.

6. *Partner learning*, by Clarence Smith, New Mexico State University.

7. *A high school visitation program*, by E. C. Boes, New Mexico State University.

8. *A comparison of direct sums of abelian semigroups*, by J. A. Anderson, Northern Arizona University.

9. *On a random walk with feedback*, by Bill Calton, New Mexico State University.

R. W. BALL, *Secretary-Treasurer*

MAY MEETING OF THE NEBRASKA SECTION

The forty-seventh annual meeting of the Nebraska Section of the MAA was held on Saturday, May 1, 1971, at Nebraska Wesleyan University, Lincoln, in conjunction with the eighty-first annual meeting of the Nebraska Academy of Sciences. Professor D. L. Skoug, Chairman of the Section, presided. There were one hundred persons in attendance of whom fifty were members of the Association.

The following officers were elected for 1971–72: Chairman, Professor Paul Haeder, University of Nebraska, Omaha; Vice-Chairman, Professor D. L. Skoug, University of Nebraska, Lincoln; Secretary-Treasurer, Professor H. M. Cox, University of Nebraska, Lincoln. Professor Barbara Buchalter of the University of Nebraska, Omaha, was elected Chairman of the Nebraska-South Dakota High School Mathematics Contest Committee. Professor J. M. Earl was recognized as Chairman of the MAA Committee on High School Mathematics Contest. Revised by-laws were adopted. Professor Garrett Birkhoff represented the officers of the Association at the meeting.

The following papers were presented:

1. *Finite topological spaces*, by G. W. Knutson, University of Nebraska, Lincoln.
2. *Extension of maps and finite topological spaces*, by J. S. Downing, University of Nebraska, Omaha.
3. *Homeomorphism groups and finite topological spaces*, by M. C. Thornton, University of Nebraska, Lincoln.
4. *Separation of nonassociates in a commutative ring by valuations*, by D. E. Brown, University of Nebraska, Lincoln.
5. *Structural properties of self-complementary graphs*, by R. A. Gibbs, Hiram Scott College.
6. *Some large composite number exercises*, by Edwin Halfar, University of Nebraska, Lincoln.
7. *Nebraska-South Dakota high school mathematics contest*, by Barbara Buchalter, University of Nebraska, Omaha, and Henry Cox, University of Nebraska, Lincoln.
8. *Preliminary report of the MAA committee on high school contests*, by James Earl, Chairman, and H. M. Cox, Executive Director.
9. *The role of modern algebra in computing*, by Garrett Birkhoff, Harvard University. An invited address.
10. *Omission and adjunction of elements in summability methods*, by S. D. Luke, Nebraska Wesleyan University.
11. *Uncountably recurrent functions*, by J. A. Eidswick, University of Nebraska, Lincoln.
12. *Innovations in the mathematics classroom*, by Sister Mary Carole Curran, Mount Marty College, Yankton, South Dakota.
13. *Teaching evaluation by questionnaire*, by M. C. Thornton, University of Nebraska, Lincoln.
14. *Game of life*, by Mark Hofmeister, Mount Marty College, Yankton, South Dakota.
15. *A reduction of generalized cubic mazes*, by Paul Payne, Hastings College.
16. *The computerized hospital*, by Mary McEnty and Sandra Krebs, Mount Marty College, Yankton, South Dakota.
17. *Euler, Pascal and the missing region*, by R. A. Gibbs, Hiram Scott College.
18. *Panel discussion—accreditation and certification in mathematics*, by Melvin George, University of Nebraska, Lincoln, Paul Haeder, University of Nebraska, Omaha, Edwin Halfar, University of Nebraska, Lincoln, and L. M. Larsen, Kearney State College.

H. M. Cox, *Secretary*

THE 1971 WILLIAM LOWELL PUTNAM MATHEMATICAL COMPETITION

The thirty-second annual William Lowell Putnam Mathematical Competition will be held on Saturday, December 4, 1971. This competition is supported by the William Lowell Putnam Intercollegiate Memorial Fund and is under the sponsorship of the Mathematical Association of America. Colleges and universities in the United States

and Canada are eligible to register undergraduates in the competition. Application forms will be mailed about October 1 to the mathematics department chairmen of the schools on the regular mailing list and to those who supervised the competition in 1970. If application material is not received by October 15, the forms can be secured by writing the director, James H. McKay, Department of Mathematics, Oakland University, Rochester, Michigan 48063. Your application should be filed with the director by October 30, 1971.

Further details are provided in the Announcement Brochure which is mailed with the registration forms. Reports of previous competitions, including examination questions, may be found in past issues of the MONTHLY. In recent years these reports have appeared in the June–July, August–September or October issue of the MONTHLY.

NEW SECTIONAL GOVERNORS OF THE ASSOCIATION

The following have been elected Governors of the Association representing the Sections indicated:

Florida	Herman Meyer, University of Miami
Illinois	Arnold Wendt, Western Illinois University
Iowa	R. V. Hogg, University of Iowa
Louisiana-Mississippi	B. E. Mitchell, Louisiana State University
Maryland-District of Columbia-Virginia	S.B. Jackson, University of Maryland
Michigan	M. S. Klamkin, Ford Motor Company, Dearborn
North Central	G. A. Heuer, Concordia College
Philadelphia	W. E. Baxter, University of Delaware
Southern California	T. M. Apostol, California Institute of Technology
Texas	C. J. Pipes, Southern Methodist University

The highest percentage of voters was 44%, occurring in the Iowa Section. The Louisiana-Mississippi Section was second with 37%.

A. B. WILLCOX, *Executive Director*

ANNOUNCEMENT OF LESTER R. FORD AWARDS

At its meeting on January 27, 1965, in Denver, Colorado, the Board of Governors authorized a number of awards, to be named after Lester R. Ford, Sr., to authors of expository articles published in the MONTHLY and the MATHEMATICS MAGAZINE. A maximum of six awards will be made annually; each award is in the amount of $100. The articles are to be selected by a subcommittee of the Committee on Publications appointed for this purpose.

The 1971 recipients of these Awards, selected by a committee consisting of Ivan Niven, Chairman, Marvin Marcus, and D. E. Richmond, were announced by President Klee at the business meeting of the Association on August 31, 1971, at Pennsylvania State University. The recipients of the Ford Awards for articles published in 1970 were the following:

J. A. Dieudonné, The Work of Nicholas Bourbaki, MONTHLY, 77 (1970) 134–145.

George Forsythe, Pitfalls in Computation, or Why a Math Book isn't Enough, MONTHLY, 77 (1970) 931–956.

P. R. Halmos, Finite-Dimensional Hilbert Spaces, MONTHLY, 77 (1970) 457–464.

Eric Langford, A Problem in Geometric Probability, MATH. MAG., 43 (1970) 237–244.

P. V. O'Neil, Ulam's Conjecture and Graph Reconstructions, MONTHLY, 77 (1970) 35–43.

Olga Taussky, Sums of Squares, MONTHLY, 77 (1970) 805–830.

HENRY L. ALDER, *Secretary*

CALENDAR OF FUTURE MEETINGS

Fifty-fifth Annual Meeting, Las Vegas, Nevada, January 19–21, 1972.

Fifty-third Summer Meeting, Dartmouth College, Hanover, New Hampshire, August 28–30, 1972.

The following is a list of the Sections of the Association with dates of future meetings so far as they have been reported to the Editorial Director.

ALLEGHENY MOUNTAIN, Pennsylvania State University, Altoona, May 5–6, 1972.

FLORIDA, Central Florida Junior College, Ocala, March 17–18, 1972.

ILLINOIS, Lake Forest College, Lake Forest, May 12–13, 1972.

INDIANA

IOWA, University of Iowa, Iowa City, April 28, 1972.

KANSAS, Washburn University, Topeka, March 24–25, 1972.

KENTUCKY

LOUISIANA-MISSISSIPPI

MARYLAND-DISTRICT OF COLUMBIA-VIRGINIA

METROPOLITAN NEW YORK

MICHIGAN

MISSOURI, Stephens College, Columbia, May 5–6, 1972.

NEBRASKA, University of Nebraska at Omaha, Omaha, April 21–22, 1972.

NEW JERSEY, Stevens Institute of Technology, Hoboken, November 13, 1971.

NORTH CENTRAL, Bemidji State College, Bemidji, Minnesota, October 23, 1971.

NORTHEASTERN, Wellesley College, Wellesley, Massachusetts, November 27, 1971.

NORTHERN CALIFORNIA

OHIO, Ashland College, Ashland, November 5–6, 1971.

OKLAHOMA-ARKANSAS, State College of Arkansas, Conway, Arkansas, March 10–11, 1972.

PACIFIC NORTHWEST, University of Washington, Seattle, June 16–17, 1972.

PHILADELPHIA, Lafayette College, Easton, November 20, 1971.

ROCKY MOUNTAIN, The Colorado School of Mines, Golden, May 5–6, 1972.

SOUTHEASTERN, Samford University, Birmingham, Alabama, March 24–25, 1972.

SOUTHERN CALIFORNIA, California Institute of Technology, Pasadena, March 11, 1972.

SOUTHWESTERN, University of New Mexico, Albuquerque, Spring 1972.

TEXAS, Southwest Texas State University, San Marcos, Apıil 1972.

UPPER NEW YORK STATE, State University College at Geneseo, New York, November 6, 1971.

WISCONSIN, Wisconsin State University, Stevens Point, April 28–29, 1972.

FUTURE MEETINGS OF OTHER ORGANIZATIONS

AMERICAN ASSOCIATION FOR THE ADVANCEMENT OF SCIENCE, Philadelphia, December 26–31, 1971.

AMERICAN MATHEMATICAL SOCIETY, Las Vegas, Nevada, January 17–20, 1972.

AMERICAN SOCIETY FOR ENGINEERING EDUCATION

ASSOCIATION FOR COMPUTING MACHINERY, Boston, Massachusetts, August 14–16, 1972.

ASSOCIATION FOR SYMBOLIC LOGIC

CENTRAL ASSOCIATION OF SCIENCE AND MATHEMATICS TEACHERS, Detroit, Michigan, November 18–20, 1971.

FIBONACCI ASSOCIATION, College of the Holy Names, Oakland, California, November 13, 1971.

INSTITUTE OF MATHEMATICAL STATISTICS

MU ALPHA THETA

NATIONAL COUNCIL OF TEACHERS OF MATHEMATICS, Chicago, Illinois, April 16–19, 1972.

OPERATIONS RESEARCH SOCIETY OF AMERICA, Disneyland Hotel, Anaheim, California, October 27–29, 1971.

PI MU EPSILON, Dartmouth College, Hanover, August 29–30, 1972.

SOCIETY FOR INDUSTRIAL AND APPLIED MATHEMATICS, University of Wisconsin, Madison. October 11–13, 1971.

GEORGE BANTA COMPANY, INC., MENASHA, WISCONSIN

THE AMERICAN MATHEMATICAL MONTHLY

(FOUNDED IN 1894 BY BENJAMIN F. FINKEL)

THE OFFICIAL JOURNAL OF

THE MATHEMATICAL ASSOCIATION OF AMERICA, INC.

VOLUME 78 NUMBER 8

CONTENTS

(Continued on inside cover)

OCTOBER 1971

NOTICE TO AUTHORS

Specialized research is usually unsuitable; see *Statement of Policy* (vol. 76, p. 2). Manuscript preparation: Please use the *Manual for Monthly Authors* (vol. 78, p. 1) and follow the format in current issues of the MONTHLY. Manuscripts should be typewritten, triple-spaced with wide margins; submit two copies and keep one for protection against loss. Each author receives 50 free reprints.

Backlog: Main Articles 7 months, Math. Notes 8 months, Research Problems 6 months, Classroom Notes 7 months, Math. Education 6 months.

EDITORIAL CORRESPONDENCE AND MAIN ARTICLES: to HARLEY FLANDERS, American Mathematical Monthly, Tel Aviv University, Ramat Aviv, Israel (see Notice, vol. 77, 1970, p. 555); NOTES, etc.: to the corresponding Associate Editor; ADVERTISING CORRESPONDENCE: to RAOUL HAILPERN, Mathematical Association of America, SUNY at Buffalo, Buffalo, N. Y. 14214; CHANGE OF ADDRESS and SUBSCRIPTIONS: to A. B. WILLCOX, Mathematical Association of America, 1225 Connecticut Ave., N.W., Washington, D. C. 20036

Annual dues for members of the Association (including a subscription to the American Mathematical Monthly) are $12.50. For nonmembers the subscription price is $18.00.

PUBLISHED BY THE ASSOCIATION at Washington, D. C., and Menasha, Wisconsin during the months of January, February, March, April, May, June–July, August–September, October, November, December.

Second-class postage paid at Washington, D. C., and additional mailing offices. Acceptance for mailing at special rate of postage provided for in the Act of February 28, 1925, embodied in Paragraph 4, Section 538, P. L. and R., authorized April 1, 1926.

PRINTED IN THE UNITED STATES OF AMERICA

MEASURE-PRESERVING TRANSFORMATIONS AND RANDOM PROCESSES

D. S. ORNSTEIN, Stanford University

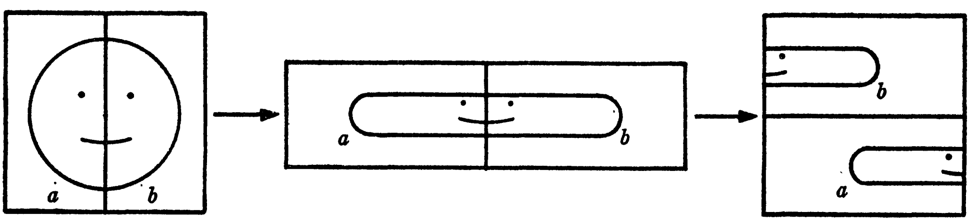
About ten years ago Kolmogorov discovered that certain ideas introduced by Shannon in information theory were relevant to the study of measure-preserving transformations. This discovery gave new life to ergodic theory and had applications to many other subjects such as probability, analysis, differential geometry, and dynamical systems.

In this article I shall make no attempt to give a survey of the above results, but I shall give a few examples which I hope will give something of the flavor of the subject.

Our results will be concerned with 1-1 invertible transformations T of a measure space X of total measure 1, where T and T^{-1} transform each measurable set onto a measurable set of the same measure. (X will always be the unit interval with Lebesgue measure or the unit square in the plane with Lebesgue measure. Both of these spaces are measure theoretically the same.)

We shall be concerned with the behavior of T only up to sets of measure 0. That is, we shall say that T_1 acting on X_1 is **isomorphic** to T_2 acting on X_2 if we can find $\overline{X}_1 \subset X_1$ and $\overline{X}_2 \subset X_2$ of measure 1 and invariant under T_1 and T_2, respectively, and if there is an invertible measure-preserving transformation ψ of $\overline{X}_1$ onto $\overline{X}_2$ such that for all x in $\overline{X}_1$ we have $\psi T_1(x) = T_2\psi(x)$.

Example 1. Let X be the unit square; T will send the point (x, y) into $(2x, \frac{1}{2}y)$ if $0 \leq x < \frac{1}{2}$ and $(2x-1, \frac{1}{2}y+\frac{1}{2})$ if $\frac{1}{2} \leq x < 1$. We can picture T as follows: We first squeeze down the unit square, and we then translate the part that has left the square back onto the part of the square that is now empty. (This nice type of picture was suggested in *Ergodic Problems of Classical Mechanics*, by V. I. Arnold and A. Avez, New York, Benjamin, 1968.)



This transformation is called the Baker's transformation.

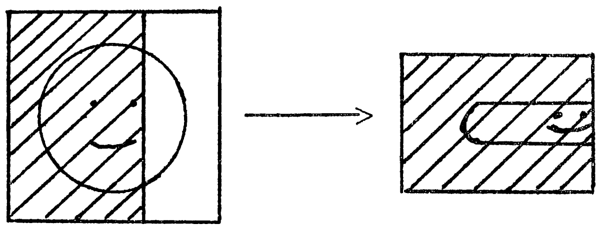
Example 2. Let X be the unit square; T will send (x, y) into $(\frac{3}{2}x, \frac{2}{3}y)$ if $0 \leq x < \frac{2}{3}$ and $(3x-2, \frac{1}{3}y+\frac{2}{3})$ if $\frac{2}{3} \leq x < 1$. We can picture T as follows:

We take the part of the square $0 \leq x < \frac{2}{3}$, $0 \leq y < 1$ and squeeze its height by

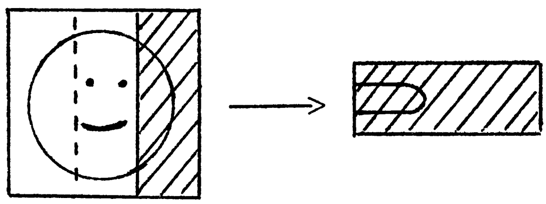
---

Donald Ornstein received his Chicago Ph.D. under Kaplansky. He spent two years at the Institute for Advanced Study and two years at the Univ. of Wisconsin before joining Stanford Univ., where he is now Professor. He spent a sabbatical year at Cornell and the Courant Inst., and his main research interests are probability and ergodic theory. *Editor.*

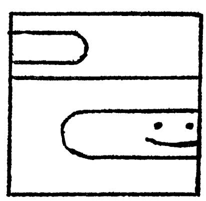
$\frac{2}{3}$ and expand its width by $\frac{3}{2}$:



We then take the part of the square $\frac{2}{3} \leqq x < 1$, $0 \leqq y < 1$ and squeeze its height by $\frac{1}{3}$ and expand its width by 3:



We then reassemble these two pieces to get



Examples 1 and 2 belong to a general class of examples which can be described as follows: X is the unit square. Divide X into k rectangles whose height is 1 and whose base is an interval of length p_i $(1 \leqq i \leqq k)$. Squeeze the height of the ith rectangle by p_i and expand its width by $1/p_i$. Now reassemble these pieces by putting the first on the bottom, the second on top of it, the third on top of that, etc. We call this transformation the **Bernoulli shift** $(p_1, \cdots, p_k)$. (The Baker's transformation is the Bernoulli shift $(\frac{1}{2}, \frac{1}{2})$.)

There is another way to describe Bernoulli shifts. Let Y be a set with k elements, and let us give the ith element measure p_i. Let Y_i, $-\infty < i < +\infty$, be copies of Y, and let X be the product of the Y_i with the product measure. Thus each point in X is a doubly infinite sequence of points in Y. Here T will act by shifting each of the above sequences. That is, T will take the sequence $\{y_i\}$ into the sequence $\{y_i'\}$, where $y_i' = y_{i+1}$. We leave it as an exercise for the reader to verify that the above definition of Bernoulli shift is the same as the first one we gave.

The main reason for giving the above form of the definition is that it shows that in a certain sense Bernoulli shifts are the simplest examples of ergodic

measure-preserving transformations (T is **ergodic** if the only sets that are invariant under T have measure 0 or 1). There is a theorem that says any invertible, ergodic, measure-preserving transformation can be obtained if we modify the above construction by taking Y to be countable and by taking some other measure invariant under the shift, instead of the product measure. Thus Bernoulli shifts are simplest in the sense that the product measure is the simplest measure invariant under the shift.

There is a third way to describe Bernoulli shifts: T is isomorphic to the Bernoulli shift $(p_1, \cdots, p_k)$ if and only if there is a partition P into sets $P_1, \cdots, P_k$ such that the measure of P_i is p_i and such that:

(1) the T^iP are independent, and

(2) the T^iP generate.

Conditions (1) and (2) mean the following: (1) means that if we take a finite number of sets of the form $T^iP_{j(i)}$ (only one for each i), then the measure of their intersection is the product of their measure; (2) means that if B is measurable, then, given ϵ, we can find n and B' in the algebra of sets generated by the $T^iP_{j(i)}$ $(-n \leqq i \leqq n,\ 1 \leqq j(i) \leqq k)$ such that the symmetric difference between B and B' is less than ϵ. (The equivalence of this third description is a routine exercise in measure theory.)

When Halmos wrote his book on ergodic theory, one of the main problems was the following: Are *all* Bernoulli shifts isomorphic? In particular, is the Bernoulli shift $(\frac{1}{2}, \frac{1}{2})$ isomorphic to the Bernoulli shift $(\frac{1}{3}, \frac{1}{3}, \frac{1}{3})$? Up to this point there were various properties of transformations that were known, and because of these a large number of transformations could be distinguished from one another (properties such as ergodicity, mixing, weak mixing). None of these properties, however, distinguished any two Bernoulli shifts (two Bernoulli shifts even induced isomorphic unitary operators).

The breakthrough in this area came when Kolmogorov introduced a new invariant called entropy, which was motivated by Shannon's work on information theory. This invariant, which I shall describe below, was easy to compute for the Bernoulli shift $(p_1, \cdots, p_k)$ and is simply $-\sum_{i=1}^{k} p_i \log p_i$. (Thus the Bernoulli shift $(\frac{1}{2}, \frac{1}{2})$ is not isomorphic to the Bernoulli shift $(\frac{1}{3}, \frac{1}{3}, \frac{1}{3})$.)

The **entropy** of a transformation is defined as follows: We first define the entropy $H(P)$ of a partition P, where the ith set has measure p_i, as $-\sum p_i \log p_i$. (If the T^iP were independent, then for all n large enough, the size of most of the atoms in the partition $\bigvee_{i=1}^{n} T^iP$ would be approximately $\frac{1}{2}^{n\cdot H(P)}$. This follows from the law of large numbers.) We now define the entropy $H(P, T)$ of P relative to T as

$$\lim_{n\to\infty} \frac{1}{n} H\left(\bigvee_{1}^{n} T^iP \right).$$

Actually the limit exists. (There is a theorem due to Shannon and McMillan that says that for all n large enough, the size of most of the atoms in $\bigvee_{i=1}^{n} T^iP$ is roughly the same and approximately equal to $\frac{1}{2}^{n\cdot H(P,T)}$.)

We now define $H(T)$ as sup $H(P, T)$, where sup is taken over all finite partitions. At first glance one would expect that $H(P, T)$ would be very large if P had a large number of atoms and that $H(T)$ would always equal ∞. To get a feeling why this is not so, note that

$$H(P, T) = \lim_{n\to\infty} \frac{1}{n} H\left(\bigvee_1^n T^i P \right)$$

should equal

$$H\left(\bigvee_1^l T^i P, T \right) = \lim_{n\to\infty} \frac{1}{n} H\left(\bigvee_1^{n+l} T^i P \right).$$

This same argument suggests that sup $H(P, T)$ should be attained if the T^iP generate. This is actually true (it would be very easy if each B were actually in some $\bigvee_{-n}^n T^iP$). It is this last fact that enables us to calculate the entropy for Bernoulli shifts.

Let us now return to the isomorphism problem for Bernoulli shifts. There are still quite a lot of Bernoulli shifts with a given entropy. Are these all the same? Is there another property yet to be discovered that will distinguish some of these? Are they all different? Mesalkin showed that the answer to the last question was no by showing that the Bernoulli shifts $(\frac{1}{4}, \frac{1}{4}, \frac{1}{4}, \frac{1}{4})$ and $(\frac{1}{2}, \frac{1}{8}, \frac{1}{8}, \frac{1}{8}, \frac{1}{8})$ were isomorphic.

Sinai proved a beautiful and deep theorem along these lines, namely: *If $H(T) = -\sum p_i \log p_i$, then we can find a partition P, whose i-th set has measure p_i, such that the T^iP are independent.* (The T^iP may not generate. If this could be shown, under the additional hypothesis that T is a Bernoulli shift, then all Bernoulli shifts of the same entropy would be isomorphic.)

We now know [**7**]:

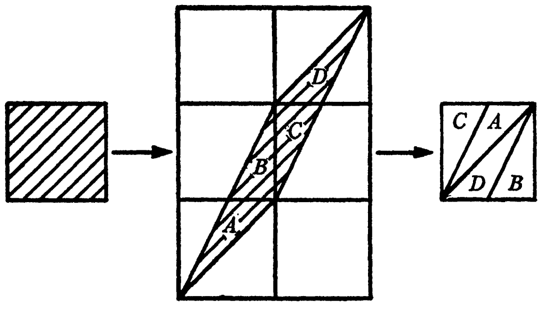
THEOREM. *All Bernoulli shifts with the same entropy are isomorphic.*

Application. There is a transformation T on the unit square in the plane such that T^2 is the Baker's transformation. (In general, Bernoulli shifts have roots of all orders.) We get the above as follows: Let T be a Bernoulli shift such that

$$H(T) = \tfrac{1}{2}(-\tfrac{1}{2}\log\tfrac{1}{2} - \tfrac{1}{2}\log\tfrac{1}{2}).$$

It is easy to see that $H(T^2) = 2H(T)$. Therefore, T^2 is isomorphic to the Baker's transformation.

Application. We shall now describe a result which does not follow from our theorem, but follows from a criterion [**8**] which comes out of the proof of that theorem. Let R be the transformation defined on the unit square as follows: $R(x, y) = (x+y, x+2y)$ (mod 1). It is easy to see that R is an invertible, measure-preserving transformation. Our criterion shows that R is (i.e., isomorphic to) a Bernoulli shift. Here is a picture of R:



R is an example of the following more general situation. Take any 2×2 matrix with integer coefficients and determinant ± 1. (In the case of R we take $\binom{1\ 1}{1\ 2}$.) Such a matrix acts on the unit square by transforming each point and then subtracting off the largest integer in the x-coordinate and the largest integer in the y-coordinate. Because the determinant is ± 1, this transformation is measure-preserving and invertible. It can be shown [1] that the transformation is ergodic if and only if the matrix has no eigenvalues that are roots of unity. In this case our criterion applies, and we can show that we get a Bernoulli shift.

Let us note that the unit square is a topological group under addition mod 1 (usually called the 2-dimensional torus) and that the transformations described above are its automorphisms. We thus have: the ergodic automorphisms of the 2-dimensional torus are Bernoulli shifts. (Adler and Weiss already knew quite a lot about such automorphisms. In particular, entropy was enough to classify them [5].)

The result extends to the 3-dimensional torus and to those n-dimensional tori whose matrices have no eigenvalues of absolute value 1. However, if the dimension is greater than 3, this result is still incomplete because there are then ergodic automorphisms that do have eigenvalues of absolute value 1. (Sinai already knew quite a lot about the automorphisms of n-dimensional tori [6].)

Returning to the general theory, we can ask how large the class of Bernoulli shifts is, or equivalently, what is the class of transformations for which entropy tells the whole story. Kolmogorov had a very general conjecture along these lines which I would now like to describe. He introduced a class of transformations (now called K-automorphisms—K-transformations would be more consistent with the terminology of this article) and conjectured that if two K-automorphisms had the same entropy, they would be isomorphic. (This would imply, of course, that they are Bernoulli shifts.)

We say that T is a K-**automorphism** if there is a finite partition P such that the T^iP generate and $\bigcap_{n=1}^{\infty} \bigvee_{j=n}^{\infty} T^iP$ is trivial. (That is: $\bigvee_{i=n}^{\infty} T^iP$ is the class of measurable sets generated by the T^iP, $n \leqq i < \infty$. The only sets which are contained in the above classes for all n have measure either 0 or 1.) A special case is when the T^iP generate and are independent.

There is a beautiful theorem, due to Sinai and Rochlin, which says that T is a K-automorphism if and only if for every finite partition Q we have that $E(Q, T) > 0$. (This is the same as saying that Q is not contained in $\bigvee_{-\infty}^{-1} T^iQ$.)

We now know that Kolmogorov's conjecture is false, [9], i.e.,

THEOREM. *There is a K-automorphism that is not a Bernoulli shift.*

The proof of this theorem is closely related to the proof of the first theorem. We need a usable criterion for when the T^iP generate a Bernoulli shift, and such a criterion comes out of the methods used in the proof of the first theorem.

One of the main problems in ergodic theory is to find out what properties tell the whole story for K-automorphisms.

One of the main sources of motivation for ergodic theory is probability theory, and I would now like to take a look at the above results from this point of view.

One of the basic and simplest objects that probability theory deals with is a stationary process. This can be thought of as a box that prints out one letter for each unit of time, where the probability of a given letter being printed out may depend on the letters printed out in the past but is independent of the time (that is, the mechanism in the box does not change).

Example 1. The box contains a roulette wheel. We spin the wheel once each unit of time and print out the result.

Example 2. The box contains a roulette wheel. We look at all possibilities for three consecutive spins and divide these into two classes. Now each time we spin the wheel we look at the last three spins and print out a 1 if they fall in the first class and a 2 if they fall in the second class.

Example 3. A teleprinter. This prints out letters where the probability of a given letter depends on what has already been printed (many possibilities will have probability 0 because they will not make sense).

The mathematical model for a process is a transformation T acting on X and a partition P of X. The reason for this is the following: We know all about the process if we know the probability of its printing out any given finite sequence. Now to each point $x \in X$ corresponds a sequence whose ith term is the atom of P containing T^ix. If we think of the measure on X as a probability measure, then the probabilities of finite sequences determine the pair P, T. (The measure of the atoms of $\bigvee_{-n}^{n} T^iP$ is determined.)

In Example 1, T is a Bernoulli shift and P is its independent generator. In Example 2 we replace P by a partition of two sets in $\bigvee_1^3 T^iP$.

Example 3 is more interesting. Our criterion is such that it seems very reasonable to assume that a teleprinter satisfies it. If this is the case, then its model will be a Bernoulli shift T and a partition P with 26 atoms (the T^iP will generate but will, of course, not be independent).

The Birkhoff ergodic theorem says that if A is any atom in $\bigvee_{-K}^{K} T^iP$, then,

except for a set of x of measure 0, the number of $1 \leqq i \leqq n$ such that $T^i x \in A$, divided by n, will tend to a limit. If T is ergodic, this limit will be the measure of A. This implies that if we look at a process long enough, each fixed string of letters will occur with a fixed frequency. If T is ergodic, this frequency will be the probability of the string.

From now on we shall consider only ergodic processes. (Any other process will be some sort of average of ergodic ones.) For an ergodic process, we can tell all about it by watching it long enough. (If we look at a large number of instances of a process and let them all print out a string of length ℓ, we can find out the probability of printing out a given string of length ℓ. If the process is ergodic, we can take any one instance of the process and look at the frequency of a particular string of length ℓ to get its probability.)

If the process P, T has 0 entropy ($H(P, T) = 0$), then P, T is deterministic in the following sense: $P \subset \bigvee_{-\infty}^{-1} T^i P$, which means that by knowing the past we can predict the next letter with probability 1, or by knowing enough of the past we can predict the next letter with arbitrarily high probability. We can even make a stronger statement. For all n, $P \subset \bigvee_{-\infty}^{-n} T^i P$, which means that we can predict a particular letter by knowing the distant past, however distant.

Now suppose T is a K-automorphism and P any finite partition such that the $T^i P$ generate. Let us call such a process a K-**process**. The K-processes are exactly those processes that contain no deterministic part or are completely non-deterministic in the following sense: If P, T is not a K-process and the $T^i P$ generate (and T is ergodic*), then there is a set S (whose measure is as close to $\frac{1}{2}$ as we want) and an N such that $S \subset \bigvee_0^N T^i P$, and for all n, S can be approximated to within 99/100 by a set in $\bigvee_{-\infty}^{-n} T^i P$. This means that we can divide the sequences of length N into two classes (each having measure close to $\frac{1}{2}$) and we predict with probability 99/100 which class the next N symbols will belong to by knowing the distant past, however distant. (All we really need to know is a finite number of terms in the distant past. In general, we need more terms in the more distant past.)

Now suppose T is a Bernoulli shift and P a finite partition. What can we say about the process P, T? Let B be a finite partition such that the $T^i B$ are independent and generate. Then the sets in P can be approximated arbitrarily well by sets in $\bigvee_{-K}^{K} T^i B$ (by choosing K large enough). This means that P, T can be approximated arbitrarily well by a box containing a roulette wheel whose distribution is the same as B, and by dividing the sequences of length $2K+1$ into classes and reporting to which class the previous $2K+1$ spins belonged. The sense in which these processes approximate P, T is the same as the sense in which flipping a coin with a little piece of dirt on one side approximates flipping a fair coin. (It can be shown that the processes P, T, where T is a Bernoulli shift, are the only processes that can be approximated by the above kind of process.) (See [10] for an exact statement.)

* The assumption is there only for simplicity. In fact, if T is not ergodic, we have an even stronger deterministic property.

Thus the processes P, T, where T is a Bernoulli shift, are those processes that can be approximated arbitrarily well by a process that can be physically constructed from a roulette wheel.

In these terms, Kolmogorov's conjecture says that if P, T is completely non-deterministic, then T is a Bernoulli shift (and the process is essentially constructable from a roulette wheel).

The fact that Kolmogorov's conjecture is false means that we have a completely non-deterministic process $\overline{P}$, $\overline{T}$ such that if T is a Bernoulli shift (and P a finite partition), then for all n large enough we can find a collection $\overline{C}$ of sequences of length n, whose measure under $\overline{P}$, $\overline{T}$ is greater than 99/100, and a collection C of sequences of length n, whose measure under P, T is greater than 99/100, and furthermore, no sequence appears in both C and $\overline{C}$. (It turns out that even more is true. There is a fixed $\delta > 0$ associated with $\overline{P}$, $\overline{T}$ such that C and $\overline{C}$ can be chosen with the additional property that any sequence in C differs from any sequence in $\overline{C}$ in more than δn places. Thus $\overline{P}$, $\overline{T}$ and P, T are a definite distance apart (see [**10**]).)

These processes are still very mysterious, and I think that it is an important and very difficult problem to get a better understanding of them.

This research was supported in part by NSF grant GP-21509

References

1. P. R. Halmos, Lectures on Ergodic Theory, J. Math. Soc. Japan, Tokyo, 1956, 54–55.

2. P. Billingsley, Ergodic Theory and Theory of Information, Wiley, New York, 1965.

3. A. N. Kolmogorov, A new metric invariant of transitive dynamical systems and automorphisms of Lebesgue spaces, Dokl. Akad. Nauk SSSR, 128 (1958) 861–864.

4. I. G. Sinai, A weak isomorphism of transformations with an invariant measure, Dokl. Akad. Nauk SSSR, 147 (1962) 797–800, and Soviet Math. Dokl., 3 (1962) 1725–1729.

5. R. L. Adler and B. Weiss, Entropy, a complete metric invariant for automorphisms of the torus, Proc. Nat. Acad. Sci. U.S.A., 57 (1967) 1573–1576.

6. I. G. Sinai, Markov partitions and U-diffeomorphisms, Funkcional. Anal. i Prilozen, 2 (1968) 64–89, and Functional Anal. Appl., 2 (1968) 61–82.

7. D. S. Ornstein, Bernoulli shifts with the same entropy are isomorphic, Advances in Math., 4 (1970), 337–352.

8. N. A. Friedman and D. S. Ornstein, An isomorphism of weak Bernoulli transformations, Advances in Math., 5 (1970) 365–394.

9. D. S. Ornstein, A Kolmogorov automorphism that is not a Bernoulli shift, to appear (Advances in Math.).

10. ———, An application of ergodic theory to probability theory, Advances in Math., to appear.

11. ———, Two Bernoulli shifts with infinite entropy are isomorphic, Advances in Math., 5 (1970) 339–348.

12. ———, Factors of Bernoulli shifts are Bernoulli shifts, Advances in Math., 5 (1970) 349–364.

13. ———, Imbedding Bernoulli shifts in flows, in "Contributions to Ergodic Theory and Probability," Lecture Notes in Mathematics Series, Springer-Verlag, Berlin, 1970, 178–218.

14. M. Smorodinsky, An exposition of Ornstein's isomorphism theorem, to appear (Advances in Math.).

15. D. S. Ornstein, Some new results in the Kolmogorov-Sinai theory of entropy and ergodic theory, invited address, Calif. Inst. of Tech., 1970, to appear in Bull. Amer. Math. Soc.

DETERMINISTIC POPULATION GROWTH MODELS

W. G. COSTELLO and H. M. TAYLOR, Cornell University

1. Introduction and Summary. Our main purpose is to present a unified and systematic derivation of the classical demographic theory of long run deterministic population growth, and this derivation occupies most of Section 2. The assumptions of this theory concern age—specific individual birth and death rates, stationary in time and homogeneous over the population. The fundamental result of the theory is that $N(t, I)$, the sub-population at time t having ages in an arbitrary age interval I, asymptotically grows exponentially at a rate r that is independent of I and of the initial conditions. More precisely, there are two cases, labeled arithmetic and non-arithmetic. In the non-arithmetic case, for some positive constant $k(I)$, depending on I and on the size and age structure of the population at time zero,

$$e^{-rt}N(t, I) \to k(I) \quad \text{as } t \to \infty. \tag{1.1}$$

An immediate corollary is that the fraction of the population having ages in an arbitrary age interval I will itself approach a limit, thus defining a long run or stable distribution of ages in the population. In the arithmetic case, (1.1) holds provided t is restricted to appropriate arithmetic subsequences. The rate r, which has been called variously, the **biotic potential,** the **Malthusian parameter,** or the **intrinsic rate of natural increase,** is considered to be of fundamental ecological significance (Cole [4]).

The precise conditions under which (1.1) holds are difficult to find in an explicit form in the literature. Most available derivations assume the existence of a stable age distribution and derive a relation similar to (1.1) from this assumption, or proceed vice versa (e.g., Lotka [25], [26]; Pielou [35]; Sharpe and Lotka [37]; Bartlett [1]; Keyfitz [18]; Karlin [17]; Gause [12]). Most derivations treat only one of two cases, discrete and continuous, or in treating both, separate them from the very start. Unfortunately these two cases do not exhaust all possibilities and, in addition, the dichotomy seems to create a need for further awkward assumptions. However, it has long been recognized (Keyfitz [18]; Karlin [17]) that the best approach to (1.1) is through a "renewal equation," and the simple, elegant treatment of renewal equations by Feller [9] affords us the opportunity of giving a general and complete proof of the asymptotically exponential population growth represented by equation (1.1), a derivation in which the breakdown is into the totally exhaustive arithmetic and non-arithmetic cases and does not appear until the final results are stated.

Prof. Taylor did his undergraduate work at Cornell and his Ph.D. work under G. Lieberman at Stanford; then he returned to Cornell, where he is now Assoc. Prof. of Operations Research and Environmental Systems Engineering. He spent an NSF Postdoctoral year at Berkeley, and his research is in applied probability and statistics.

William Costello received his B.S. from Tufts in 1959 and spent the next 10 years in the U.S. Navy before resigning with the rank of Lt. Commander. He is currently a graduate student at Cornell in Operations Research. *Editor.*

The original development of an age specific population growth theory is due to Lotka [25–30] in a differentiable (hence nonarithmetic) setting. If $b(u)du$ are the number of progeny of an individual during his infinitesimal age span $[u, u+du]$, and if $l(u)$ is the fraction of individuals in the population that survive to at least age u, then Lotka shows that the growth rate r in equation (1.1) is the unique positive solution to

$$g(r) = \int_0^\infty e^{-rx}l(x)b(x)dx = 1, \tag{1.2}$$

provided $g(0)>1$. We define a **net maternity function**

$$m(u) = b(u) \times l(u)$$

as a birth rate adjusted for the death of some fraction of the population, and a **cumulative net maternity function**

$$M(u) = \int_0^u m(x)dx$$

as the total number of progeny during the interval $(0, u]$ of an individual born at time zero. In Lotka's differentiable case, M is continuous. We take M as our point of departure and assume it continuous from the right, non-decreasing, and bounded, with $M(0)=0$ and a total variation $M(\infty)>1$. In this manner we are able to unify up to some point, the non-arithmetic case, which includes Lotka's formulation, and the arithmetic case, related to the discrete time formulation of Thompson [39], Lewis [23], Leslie [22], [21], and others.

All these formulations, in addition to ignoring the stochastic aspects of the problem, also disregard the effects of the environment upon the population, for example in food supply or space limitations, as well as any interaction with other species. It is also assumed that the population is completely homogeneous in its structure, distribution, and growth. Attempts have been made to include the effects of environmental resistance to growth, generalizing (1.1) to something of the form $N(t) \sim k \exp\{\int_0^t rf(N(t))dt\}$, where $f(N)$ includes the environmental resistance, depending upon the current population size. An historically prominent function is the **logistic,** attributable to Verhulst [41], in which the resistance is inversely proportional to the remaining capacity for growth in the environment. In the case of the human population, there are no signs that environmental resistance has yet played any significant role, while the current growth in numbers of humanity appears unbounded, and indeed growing at a faster than exponential pattern (e.g., see Dorn in Bresler [6]). This has invoked increasing concern among the public at large, not only because of the possible social consequences or psychological effects, but also for deleterious side effects such as environmental deterioration.

Nevertheless, one might question the ecological relevance of a model which inevitably leads to the unbounded growth represented by equation (1.1). As Gause [12] quotes his teacher, Vernadsky, "For every species or race there is

a maximal number of individuals which can never be surpassed. This maximal number is reached when the given species occupies entirely the earth's surface, with a maximal density of its occupation." But even so, equation (1.1) is regarded as quite important by ecologists. It is interpreted (Cole, [4]) as representing a potential rate of growth, characteristic of a species, and describing the "ideal" growth of that species unlimited by environmental restrictions. We discuss the ecological significance of the model and some of its possible uses in helping plan human population growth in Section 3.

Throughout, a variety of potential research areas are indicated.

2. **Age structured growth.** For each $u \geqq 0$, let $M(u)$ be the cumulative number of (female) progeny an arbitrary (female) individual will contribute up to age u. Of course M, called the **cumulative net maternity function,** is non-decreasing, and we assume M to be bounded, continuous from the right, with $M(0)=0$, and $M(\infty)>1$. (Warning: M is also known as the **net fertility function.** However, "fertility" has a variety of conflicting meanings, and we have tried to avoid using the term where possible. See the Note on Terminology in *Public Health and Population Change*, M. C. Sheps and J. C. Ridley, eds., Univ. Pittsburgh Press, 1965.) Write $M[\cdot]$ for the Lebesgue-Stieltjes measure on the Borel real line that is induced by M. Thus, for an interval $I=(a, b]$, we have $M[I]=M(b)-M(a)$, and $M[dx]=dM(x)$. Without further mention, we shall always so identify non-decreasing right continuous functions with their related measures. For a bisexual population, the theory considers explicitly only the female members of the population, assuming that males are present in an identical age distribution, so that one extrapolates results directly to the entire population using the fixed sex ratio. Thus $M(u)$ is the number of female offspring a female contributes up to age u, $N(t)$ is the number of females in the population at time t, and so on. Biological and social categorizations other than age are disregarded in what we present here, although extensions are possible (Goodman [14]; Keyfitz [18]). Note finally that the cumulative number of progeny of an individual depends in our model only on the individual's age and not otherwise on time, so that our model is stationary in this sense.

We do not hesitate to allow M to assign a fractional number of progeny to an age interval I. Indeed, Lotka concentrated his attention on the case of a differentiable M. Presumably, fractional progeny are to be interpreted as mean values per individual in a randomly evolving but large population. Incidentally, the study of stochastic versions of these age-specific population growth models is just beginning (Jagers [16]). With a similar interpretation of fractional values, let $l(t)$, called the **survivorship function,** be the fraction of individuals in a cohort that survive to age t. By definition, M has been reduced to reflect survivorship; the number of births in an infinitesimal age interval du for a parent that has survived to age u is $M[du]/l(u)$. We shall assume that the total progeny, given that a parent lives throughout the childbearing years, is finite, i.e.,

$$\int_0^\infty l(u)^{-1}M[du] < \infty. \tag{2.1}$$

Let $B(t)$ be the cumulative number of new births in the population during the half open time interval $(0, t]$, and let $B_0(t) \leqq B(t)$ be those births from individuals already in the population at time zero. If $N(t, u)$ is the number of individuals in the population at time t of age not exceeding u, then the $N[0, du]$ individuals of age $(u, u+du]$ in the population at time zero are the result of $l(u)^{-1}N[0, du]$ individuals being born infinitesimally near time $-u$, each of which produces $M(t+u)-M(u)$ offspring in the time interval $(0, t]$. Summing over u, we get

$$B_0(t) = \int_0^\infty \{M(t+u) - M(u)\} l(u)^{-1} N[0, du], \tag{2.2}$$

where we agree that the integrand is zero whenever $l(u)=0$. Then

$$\int_0^\infty l(y)^{-1} M[dy] \geqq \int_u^{u+t} l(y)^{-1} M[dy] \geqq l^{-1}(u)\{M(u+t) - M(u)\},$$

which with (2.1) implies

$$B_0(\infty) = \sup_t B_0(t) < \infty, \tag{2.3}$$

provided, of course, $N(0) = \int N[0, du] < \infty$.

With $B(t)$ the total births and $B_0(t)$ the births from parents present in the population at time zero, the difference $B(t)-B_0(t)$ counts the progeny of individuals themselves born after time zero, say at time $x \leqq t$, and of age $u=t-x$ at time t. Summing over x, we get

$$\begin{aligned} B(t) &= B_0(t) + \int_0^t M(t-x) B[dx] \\ &= B_0(t) + \int_0^t B(t-x) M[dx] \quad \text{for } t \geqq 0. \end{aligned} \tag{2.4}$$

We extend the definition of all functions mentioned in (2.4) by agreeing that they vanish for negative arguments. With the extension, the integration may be performed over the whole real line with the same result.

The analysis of the asymptotic population growth centers about equation (2.4), called a "renewal equation" in probability theory, and we shall present without proof three theorems concerning such equations. Our reference is Feller [9], especially Chapter XI and Section 6 of Chapter VI, where proofs and references to the original works may be found.

All functions introduced in this discussion are assumed to vanish on $(-\infty, 0)$. For non-negative, right continuous, non-decreasing functions A and B, with B bounded, write $A*B$ for the convolution defined by

$$A * B(x) = \int_0^x B(x-y) A[dy]. \tag{2.5}$$

Then we may rewrite (2.4) as:

$$B(t) = B_0(t) + B * M(t), \qquad \text{for} \quad t \geqq 0,$$

or even $B = B_0 + B * M$.

A **renewal process** is a sequence $\{T_k\}$ of independent and identically distributed positive random variables, representing the lifetimes of some "units." The first unit is placed in operation at time zero; it fails at time T_1 and is replaced by a new unit which then fails at time T_1+T_2, and so on, thus motivating the name "renewal process." The time of the $n+1$st renewal is $S_n = T_1 + \cdots + T_n$, where, arbitrarily, we count the first renewal as occurring at time $S_0 = 0$.

Let F be the common probability distribution function of the T_k's, subject to the convention $F(0) = 0$. Define

$$F^{(0)}(x) = \begin{cases} 1 & x \geqq 0 \\ 0 & x < 0, \end{cases}$$

and for $n \geqq 1$, $F^{(n)} = F^{(n-1)} * F$. For any $t \geqq 0$, the $n+1$st renewal occurs in $[0, t]$ if and only if $S_n \leqq t$. By counting renewals and taking expectations, we find that the expected number of renewals in $[0, t]$, denoted by $U(t)$, is given by:

$$U(t) = \sum_{n=0}^{\infty} P_n\{S_n \leqq t\} = \sum_{0}^{\infty} F^{(n)}(t). \tag{2.6}$$

Argued formally, if we separate the first term in the sum and factor one convolution outside the summation, we see that U satisfies the renewal equation

$$U(t) = 1 + F * U(t), \qquad t \geqq 0. \tag{2.7}$$

Argued probabilistically, the expected number of renewals in $[0, t]$ is one plus the expected number in $(0, t]$, the remainder of the interval. This interval contains renewals only if $T_1 \leqq t$; given that $T_1 = y \leqq t$, the expected number of renewals in $(0, t]$ is $U(t-y)$. Equation (2.7) results from summing over y.

The precise statement is given in Theorem 1, from Feller [**9**, p. 183].

THEOREM 1. *Define $U = \sum_0^\infty F^{(n)}$. Then $U(x) < \infty$ for all x. If z is a bounded function vanishing for $x < 0$, then the convolution $Z = U * z$ represents a solution of the renewal equation $Z = z + F * Z$. There exists no other solution vanishing on $(-\infty, 0)$ and bounded on finite intervals.*

Note in particular for $z = F^{(0)}$, that $U = \sum_0^\infty F^{(n)}$ is the unique solution to equation (2.7) vanishing on $(-\infty, 0)$ and bounded on finite intervals.

The subsequent analysis suffers from a necessary division into two cases.

DEFINITION. A distribution F is called **arithmetic** if it is concentrated on a set of points of the form $0, \pm\lambda, \pm 2\lambda, \cdots$. The largest λ with this property is called the **span** of F.

Let $\mu = \int_0^\infty xF[dx] = \int_0^\infty [1 - F(x)]dx$. When $\mu = \infty$, interpret μ^{-1} as 0. The basic limit theorem of renewal theory is Theorem 2, again quoted from Feller [**9**, p. 347].

THEOREM 2. *If F is not arithmetic, then*

$$U(t) - U(t-h) \to h/\mu \quad \textit{as} \quad t \to \infty \tag{2.8}$$

for each fixed $h>0$. If F is arithmetic, the same is true when h is a multiple of the span λ.

This basic limit implies others as shown in our last theorem from Feller.

DEFINITION. Let z be a function vanishing on $(-\infty, 0)$. For fixed $h>0$ denote by $m^*(n)$ and $m_*(n)$, respectively, the supremum and infimum of z in the interval $(n-1)h \leqq x \leqq nh$. The function z is called **directly Riemann integrable** if $\sigma^* = h\sum m^*(n)$ and $\sigma_* = h\sum m_*(n)$ converge absolutely and if $\sigma^* - \sigma_* \leqq \epsilon$ for an arbitrary $\epsilon>0$, provided h is sufficiently small.

THEOREM 3 (Feller [9, p. 349]). *Let z be directly Riemann integrable, vanishing on $(-\infty, 0)$, and let $Z = U * z$ be the solution to $Z = z + F * Z$. Then*

$$\lim_{t\to\infty} Z(t) = \mu^{-1}\int_0^\infty z(y)dy$$

when F is non-arithmetic, and

$$\lim_{n\to\infty} Z(t+n\lambda) = (\lambda/\mu)\sum_k z(t+k\lambda)$$

when F is arithmetic with span λ.

Before these, our main tools, are applicable, one final change has to be made. The distribution function F in equation (2.7) and its followers is a probability distribution having total mass one, while M, its counterpart in equation (2.4), has a mass exceeding one by assumption. Note that $g(\rho) = \int_0^\infty e^{-\rho x}M[dx]$ is continuous and strictly decreasing in ρ, $g(0) = M(\infty) > 1$, and $g(\rho) \to 0$ as $\rho \to \infty$. It follows that we may define

$$M^\#[dx] = e^{-rx}M[dx], \tag{2.9}$$

where $r>0$ is the unique real root of $g(r) = 1$. Equation (2.4) becomes

$$e^{-rt}B(t) = e^{-rt}B_0(t) + \int_0^t e^{-r(t-x)}B(t-x)M^\#[dx], \qquad t \geqq 0, \tag{2.10}$$

and we immediately conclude from Theorem 1 that

$$B(t) = \int_0^t e^{rx}B_0(t-x)U[dx] < \infty, \tag{2.11}$$

where

$$U(t) = \sum_{n=0}^\infty M^{\#(n)}(t) \tag{2.12}$$

is a solution to (2.4) and is unique among solutions bounded in finite intervals and vanishing on $(-\infty, 0)$.

Of course, B_0 is monotonic, and equation (2.3) states that B_0 is bounded. Now it is a straightforward exercise in elementary analysis to show that $e^{-rt}B_0(t)$ is directly Riemann integrable. Thus, it follows from Lemma 1 and Theorem 3 that

$$e^{-rt}B(t) \to K \quad \text{as} \quad t \to \infty, \tag{2.13}$$

where

$$K = \frac{\int_0^\infty e^{-rx}B_0(x)dx}{\int_0^\infty xe^{-rx}M[dx]} \tag{2.14}$$

when M is non-arithmetic. When M is arithmetic with span λ,

$$e^{-r(t+n\lambda)}B(t+n\lambda) \to C_t \quad \text{as} \quad n \to \infty, \tag{2.13$'$}$$

where

$$C_t = \frac{\lambda \sum_k e^{-r(t+k\lambda)}B_0(t+k\lambda)}{\int_0^\infty xe^{-rx}M[dx]}. \tag{2.14$'$}$$

We remark that $K=0$ when the denominator in (2.14) is infinite, and similarly for C_t in (2.14′). Note also that, since B is monotonic,

$$\sup_t e^{-rt}B(t) \leqq \max\{B(\tau), K+1\} < \infty, \tag{2.15}$$

where τ is chosen, according to (2.13), such that $t \geqq \tau$ implies $e^{-rt}B(t) \leqq K+1$.

At this stage we've shown that asymptotically for large t, the cumulative births grow exponentially at rate r. It follows that asymptotically the whole population, and every age group within it, will grow at the same rate. To prove this, let $N(t, u)$ be the number of people in the population at time t of an age not exceeding u. For $t \geqq u$, these individuals were all born on or after time zero, so that

$$N(t, u) = \int_{t-u}^t l(t-x)B[dx] = \int_0^t l_u(t-x)B[dx], \qquad t \geqq u,$$

where

$$l_u(v) = \begin{cases} l(v) & 0 \leqq v \leqq u \\ 0 & u < v. \end{cases}$$

Let $m_u(x)=1-l_u(x)$, a bounded, nondecreasing, right continuous function, and integrate by parts to get

$$B(t)-N(t,u)=\int_0^\infty B(t-x)m_u[dx]$$

or

$$e^{-rt}B(t)-e^{-rt}N(u,t)=\int_0^\infty e^{-rx}e^{-r(t-x)}B(t-x)m_u[dx].$$

By (2.13), (2.15), and dominated convergence, the right hand side converges to $K\int_0^\infty e^{-rx}m_u[dx]$ in the non-arithmetic case, while the first term on the left converges to K. It follows that

$$\begin{aligned}\lim_{t\to\infty} e^{-rt}N(t,u) &= K\left\{1-\int_0^\infty e^{-rx}m_u[dx]\right\}\\ &= rK\int_0^u e^{-rx}l(x)dx.\end{aligned} \tag{2.16}$$

Each class then grows asymptotically at the same rate r. It's no harder to show that (2.16) holds for $u=\infty$, so that the population as a whole shares the same behavior. The limiting fraction of the population of an age not exceeding u is

$$\lim_{t\to\infty}\frac{N(t,u)}{N(t)}=A(u), \tag{2.17}$$

with

$$A(u)=\frac{\int_0^u e^{-rx}l(x)dx}{\int_0^\infty e^{-rx}l(x)dx}, \tag{2.18}$$

the stable age distribution which plays such a central role in Lotka's original development. In the case of arithmetic M with span λ we get

$$\lim_{n\to\infty} e^{-r(t+n\lambda)}N(t+n\lambda,u)=rC_t\int_0^u e^{-rx}l(x)dx \tag{2.16$'$}$$

and

$$\lim_{n\to\infty}\frac{N(t+n\lambda,u)}{N(t+n\lambda)}=A(u), \tag{2.17$'$}$$

where $A(\cdot)$ remains as in (2.18).

The complete time dependent expression for $N(t,u)$ is

$$N(t, u) = \int_0^{(u-t)^+} [l(t+v)/l(v)] \cdot N[0, dv] + \int_{(t-u)^+}^{t} l(t-v)B[dv], \tag{2.19}$$

which, in conjunction with equation (2.4), may be used to investigate numerically time dependent solutions. Numerical Fourier transforms (Goertzel, [**13**]) probably provide the most efficient means for arriving at answers on digital computers. The bound on $e^{-rt}B(t)$ expressed by equation (2.15) and integration by parts show that the Laplace-Stieltjes transform

$$B^*(\theta) = \int_0^\infty e^{-\theta x} B[dx] \tag{2.20}$$

converges for $\theta > r$. If we similarly define

$$B_0^*(\theta) = \int_0^\infty e^{-\theta x} B_0[dx] \tag{2.21}$$

and

$$M^*(\theta) = \int_0^\infty e^{-\theta x} M[dx] \tag{2.22}$$

and take transforms in equation (2.15), we get

$$B^*(\theta) = B_0^*(\theta) + B^*(\theta)M^*(\theta), \quad \text{for } \theta > r$$

or

$$B^*(\theta) = B_0^*(\theta)/[1 - M^*(\theta)], \quad \text{for } \theta > r,$$

which may be inverted to arrive at $B(t)$. A special case arises when the population begins with a single individual of age zero at time zero, which from (2.2) gives $B_0(t) = M(t)$ and equation (2.22) becomes

$$B^*(\theta) = M^*(\theta)/[1 - M^*(\theta)] = \sum_1^\infty M^*(\theta)^n \tag{2.23}$$

for $\theta > r$.

As mentioned earlier, the use of modern Fourier transform techniques in conjunction with equation (2.22) or (2.23) probably represents one of the most efficient ways of computing time dependent solutions. Interestingly enough, the special discrete time version of these transforms was proposed as an efficient numerical technique by Thompson [**39**], who attributes the idea to H. E. Soper. Let us suppose that M places mass m_t on point t for $t = 1, 2, \cdots$, and let

$$m^*(s) = \sum_{t=1}^\infty m_t s^t = M^*(\theta) \qquad \text{for } s = e^{-\theta}. \tag{2.24}$$

As an example (Cole, [4]), "two offspring at one year and two at the second year" yields

$$m^*(s) = 2s + 2s^2. \tag{2.25}$$

Then b_t, the number of births at time t, is the coefficient of s^t in

$$b^*(s) = B^*(\theta) = \sum_{t=1}^{\infty} b_t s^t, \qquad s = e^{-\theta}. \tag{2.26}$$

If we suppose the population begins with one individual at time zero, equation (2.23) yields

$$b^*(s) = m^*(s)/[1 - m^*(s)] = [1 - m^*(s)]^{-1} - 1. \tag{2.27}$$

To continue Cole's example in which $m^*(s) = 2s+2s^2$, we get

$$[1 - 2s - 2s^2]^{-1} = 1 + 2s + 6s^2 + 16s^3 + 44s^4 + 120s^5 + 328s^6 + \cdots, \tag{2.28}$$

showing that the one original female gives rise to 328 female offspring in the sixth year. We leave to the reader the interesting combinatorial exercise of showing directly that equation (2.28) has this interpretation.

3. Some uses of the model. It is interesting that the unbounded growth model of Section 2 is of importance even when applied to clearly limited populations. Equation (2.16), which expresses the exponential growth $N(t) \sim e^{rt}$, is then interpreted as a potential for growth. The potential growth rate r, determined by $\int e^{rx} M[dx] = 1$, is regarded as a more or less fixed species characteristic, independent of environmental factors, such as limited food, predation, and so on, that limit the realized growth. This conceptualization of a system balanced between a potential ability to grow and environmental resistance to growth has an early history in ecology. Darwin [5, pp. 56–57] says, "All we can do, is to keep steadily in mind that each organic being is striving to increase in a geometrical ratio; that each at some period of its life, during some season of the year, during each generation, or at intervals, has to struggle for life and to suffer great destruction." Vernadsky [42, p. 37] says " . . . the multiplication of all organisms can be expressed by geometric progression. This can be evaluated by a uniform formula:

$$2^{bt} = N_t.$$

. . . Parameter b is characteristic for every kind of living being. . . . It can be regarded as empirically established that the process of multiplication is retarded in its manifestation only by external forces; . . . "

Chapman [2], [3] labels Vernadsky's b or, equivalently, our r the "partial potential," a name refined by Gause [12, p. 33] to "biotic potential." Gause says, "The biotic potential represents the potential rate of increase of the species under given conditions. It is realized if there are no restrictions of food, no toxic waste products, etc." Cole [4] follows Lotka and terms r "the intrinsic rate of

natural increase," which he says " . . . must be regarded as a quantity of fundamental ecological significance . . . to be interpreted as the rate of true compound interest at which a population would grow if nothing impeded its growth and if the age-specific birth and death rates were to remain constant."

We spent this time clarifying the ecological meaning of r in order to introduce a class of interesting problems, first mentioned by Lotka [**27**]. Roughly speaking, the idea is this: Age specific birth patterns must, to some extent, be subject to natural selection, and one might suppose that this selection would be in the direction of increasing r. Thus Lotka poses the mathematical problem of maximizing r subject to $\int e^{-rx}M[dx]=1$, where M varies over some class determined by survivorship, energy considerations, and so on. Lotka [**27**, p. 128] says, "In point of fact, [the birth rate] is undoubtedly for most species of organisms, very elastic (much more so than the survival factor), and capable of adapting itself to varying circumstances. This is especially so in the case of man, who exhibits in particularly high degree the rather astonishing phenomenon of a portion of matter whose growth is at least partially under the control of a *will* in some manner associated with it. But in the organic world at large also, there is presumably at least some tendency to the adjustment of the procreation factor so to take place as to make the rate of increase r a maximum under the existing conditions. Too high a procreation factor would lead to excessive sacrifices of progeny that could not be raised to maturity, and would increase the death rate more than the birth rate. On the other hand, too small a procreation factor would obviously fail to give the maximum attainable rate of increase. Somewhere between the two extremes, a certain optimum procreation factor will make r a maximum."

This evolutionary behavior proposed by Lotka is now termed **r-selection.** The conceptualization of r-selection has been further developed by Fisher [**10**], and by many others since. Recently, for example, MacArthur and Wilson [**31**] have indicated some of the types of situations and environments in which r-selection might predominate over selection or other factors, and Gadgil and Bossert [**11**] have defined a specific model of r-selection and have investigated their model numerically. Except for Gadgil and Bossert, most of the research has been biologically oriented, and the problem seems ripe for further mathematical study.

Lotka notes that man's procreation is partially under the control of a will, which brings us back to one of the earliest proponents of "biotic potential" versus "environmental resistance." Malthus [**32**], as everyone knows, proposed that population grows exponentially while resources grow arithmetically, and that eventually the former would outstrip the latter. Sadler [**36**, Book III] laboriously computed at what ages persons would have to marry and how often they would have to reproduce to give some of the rates of doubling postulated by Malthus. With the current interest in planned population growth, similar to Sadler, we can use the equation $\int e^{-rx}M[dx]=1$ to see how proposed birth patterns M would reflect themselves in long run growth rates r.

Cole performed similar calculations which indicate that each litter (birth)

contributes less to potential population growth than the one preceding it; that, as the age at first reproduction increases, the effect of increased family size becomes less. His results therefore emphasize the importance in human population growth of age at marriage and age at first reproduction in contrast to total family size. Slobodkin [38] extends this to comment that in the case of man: "the number of births per female lifetime is in general of less significance in determining the reproductive potential of the population than is the age at initial reproduction."

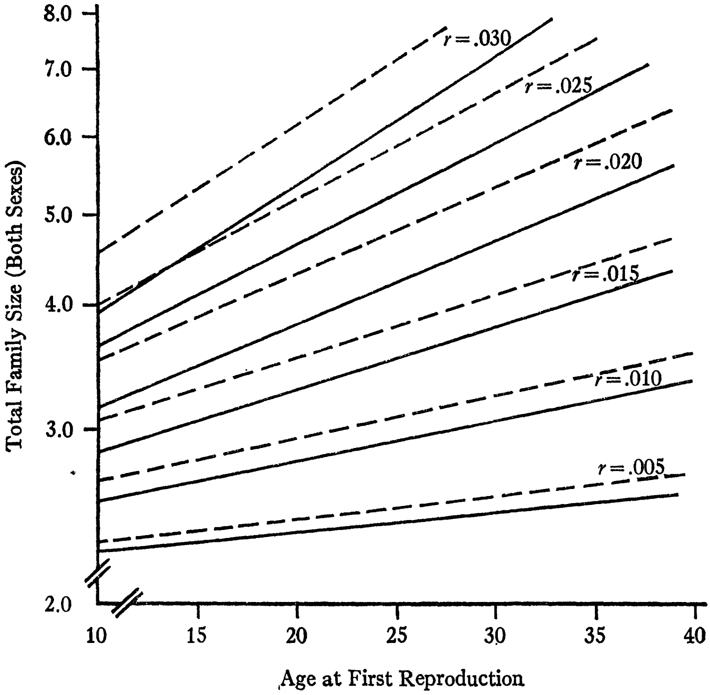


FIG. 1. Plot of Total Family Size against Age at First Reproduction for six values of intrinsic growth rate r using a gamma density for net maternity with $\beta=.10$ (dashed lines) and $\beta=.143$ (solid lines), and $A=1$.

Intrigued by the potential implication of these results in population planning, we attempted similar calculations. Our objective was to search for qualitative results concerning the relative importance of total family size versus age at first reproduction in determining population growth rates. Our results are summarized in Figure 1. We were impressed by the strong interaction of the life history parameters which determine the growth pattern. We found that the question of whether one parameter is of more or less significance than the other may not be as simple as Cole and Slobodkin suggest, but may depend very strongly on the values of the parameters themselves.

In planning our calculations we felt that the form of M, the net maternity

function, would be of great importance. Cole [4] used a uniform survivorship, $l(u)=1$ for $\alpha \leqq u \leqq \omega$, and 0 otherwise, where α is the age at first reproduction, and ω is the age at last reproduction. He used a discrete age-specific maternity function describing spikes of constant size (litter size) b at intervals of one time unit from age α to age ω. Keyfitz [18] describes calculations in which the net maternity function m is a normal distribution (Lotka [30]), a gamma distribution (Wicksell [44]), and an exponential distribution (Hadwiger [15]). While there is no theoretical justification for any of these, they all share some of the characteristics of human life history data. Keyfitz reported inconclusive results in comparing the statistical fit of different maternity functions to population data, due to the insufficiency of the latter. Our calculations were done with a variety of gamma distributions, a form which is suggested by data on human populations (Lewontin [24]), and one which possesses some analytical tractability. We used the net maternity density:

$$m(x) = \begin{cases} \dfrac{s\beta}{\Gamma(A+1)} [\beta(x-\alpha)]^A e^{-\beta(x-\alpha)} & \text{for } x \geqq \alpha, \\ 0 \quad \text{for } x < \alpha, \end{cases}$$

where

$s=\int_0^\infty m(x)dx$ is the family size (females),
$x=\alpha+A/\beta$ is the age of maximum net maternity,
$x+1/\beta$ is the mean, and
$\alpha+\sqrt{(A+1)}/\beta$ is the standard deviation of the net maternity function.

The quantities β and A determine the shape of the distribution, $1/\beta$ is the mean length of child-bearing years after the year of peak maternity, determined by A/β, while the spread of childbearing, represented by the standard deviation, depends on $(A+1)^{1/2}/\beta$. These considerations lead to approximate ranges of

$$.15 \leqq \beta \leqq .40, \quad \text{and} \quad .15 \leqq A \leqq 3.0.$$

Taking the Laplace-Stieltjes transform of $m(x)$ and setting it equal to 1 as in (2.21) will yield $(1+r/\beta)^{A+1}=se^{-r\alpha}$. If we let $S=2s$ be the total family size, including both sexes assuming they appear in equal numbers in the population, we get $2(1+r/\beta)^{A+1}=Se^{-r\alpha}$. Using this relationship, we plot the logarithm of the total family size S (both sexes) against the age at first reproduction α for different values of r, β, and A. This appears in Figure 1.

Consider ranges of S, α, and r applicable to human populations. Then Figure 1 shows that increasing the age at first reproduction yields diminishing returns in reduction of r as r comes closer to 0. Of course, in the limit, to achieve $r=0$ requires a family size of exactly 1 (female) which exactly reproduces the female parent, and the age at first reproduction matters not.

From Fig. 1, for example, if $s=4$ (2 male, 2 female), $\alpha=16$, $A=1$, $\beta=.10$, and

$r = .020$, and one wished to reduce r to, say .015, then one could use one or a combination of the following alternatives:

(1) increase α from 16 to 28;

(2) increase α from 16 to 33, increasing β from .1 to .143;

(3) reduce S from 4 to 3.3; or

(4) reduce S from 4 to 3.1, increasing β from .1 to .143.

Even this relatively small change in r from .020 to .015 requires changes in life histories which are easier discussed than accomplished. Furthermore, one must recognize the interdependence of these parameters in reality. We suspect that it would be difficult to cause a change in, say, α without causing possibly unforeseen changes in other parameters. And, of course, r is a long term rate. The short term consequences of any proposed change would have to be carefully considered.

This paper is an outgrowth of a Workshop in Population Ecology, sponsored by the Center for Environmental Quality Management, Cornell University, 1969–1970 and was supported in part by PHS Grant 1 TO1 ES 00130-03. It was presented at the 1970 Summer meeting of the MAA in Laramie.

References

1. M. S. Bartlet, Stochastic Population Models, Methuen, London, 1960.

2. R. N. Chapman, Animal Ecology, McGraw-Hill, New York, 1938.

3. ———, The quantitative analysis of environmental factors, Ecology, 9 (1928), 111–122.

4. L. C. Cole, The population consequences of life history phenomena, Quart. Review of Biology, 29 (1954) 103–137.

5. C. Darwin, The Origin of Species by Means of Natural Selection (1895) [1899 edition, London, John Murray].

6. H. F. Dorn, "World Population Growth: an International Dilemma" in Bresler, Jack B. (Editor), Human Ecology, Addison-Wesley, Reading, Mass., 1966.

7. L. I. Dublin and A. J. Lotka, On the true rate of natural increase as exemplified by the population of the United States, 1920, J. Amer. Statist. Assoc., 20 (1925) 3–5–339.

8. O. Dunkel, Solutions of a probability difference equation, this MONTHLY, 32 (1925) 354–370.

9. William Feller, An Introduction to Probability Theory and Its Applications, Volume II, Wiley, New York, 1966.

10. R. A. Fisher, The Genetical Theory of Natural Selection, Oxford Univ. Press, London, 1930.

11. Madhav Gadgil and W. H. Bossert, Life historical consequences of natural selection, The American Naturalist, 104 (1970) 1–24.

12. G. F. Gause, The Struggle for Existence, Williams and Wilkins, Baltimore, 1934.

13. G. Goertzel, Fourier Analysis. Mathematical Methods for Digital Computers, Volume I (Ralston and Wilf, eds.), Wiley, New York, 1960.

14. L. A. Goodman, The analysis of population growth when the birth and death rates depend on several factors, Biometrics, 25 (1969) 659–681.

15. H. Hadwiger, Über die Integralgleichungen der Bevölkerungstheorie, Mitteilungen der Vereinigung schweizerischer Versicherungsmathematiker, 38 (1939) 1–14.

16. P. Jagers, A general stochastic model for population development, Skand. Aktuarietidskrift, 52 (1969) 84–103.

17. Samuel Karlin, A First Course in Stochastic Processes, Academic Press, New York, 1966.

18. Nathan Keyfitz, Introduction to the Mathematics of Population, Addison-Wesley, Reading, Mass., 1968.

19. V. A. Kostitzin, Mathematical Biology, Harrap, London, 1939.

20. P. H. Leslie and J. Park, The intrinsic rate of natural increase of Tribolium castaneum Herbat, Ecology, 30 (1949) 469–477.

21. P. H. Leslie, Some further notes on the use of matrices in population mathematics, Biometrika, 35 (1948) 213–245.

22. ———, The use of matrices in certain population mathematics, Biometrika, 33 (1945) 183–212.

23. E. G. Lewis, On the generation and growth of a population, Sankhyā, 6 (1942) 93–96.

24. R. C. Lewontin, Selection for colonizing ability, The Genetics of Colonizing Species (Baker and Stebbins, eds.), Academic Press, New York, 1965.

25. A. J. Lotka, Relation between birth rates and death rates, Science, 26 (1907) 21–22.

26. ———, Studies on the mode of growth of material aggregates, Amer. J. Science, 24 (1907) 199–216.

27. ———, Elements of Physical Biology, Williams and Wilkins, Baltimore, 1925.

28. ———, The structure of a growing population, Human Biol., 3 (1931) 459–493.

29. ———, "Théorie analytique des associations biologiques," Part I "Principes," Actualités Scientifiques et Industrielles, No. 187 Hermann, Paris, 1934.

30. ———, "Théorie analytique des associations biologiques," Part II "Analyse démographique avec application particulière à l'espèce humaine," Actualités Scientifiques et Industrielles, No. 780 Hermann, Paris, 1939.

31. R. H. MacArthur and E. O. Wilson, The Theory of Island Biogeography, Princeton University Press, New Jersey, 1967.

32. T. R. Malthus, An essay on the principle of population as it affects the future improvement of society, with remarks on the speculations of Mr. Godwin, M. Condorcet, and other writers, Printed for J. Johnson in St. Paul's Churchyard, London.

33. P. A. P. Moran, Some remarks on animal population dynamics, Biometrics, 6 (1950) 250.

34. A. J. Nicholson, An outline of the dynamics of animal populations, Austr. J. Zoology, 2 (1954) 9.

35. E. C. Pielou, An Introduction to Mathematical Ecology, Wiley-Interscience, New York, 1969.

36. M. T. Sadler, The Law of Population: A Treatise, in Six Books; in Disproof of the Superfecundity of Human Beings, and Developing the Real Principle of their Increase, Murray, London, 1830.

37. F. R. Sharpe and A. J. Lotka, A problem in age distribution, Phil. Mag., 21 (1911) 435–438.

38. L. B. Slobodkin, Growth and Regulation of Animal Populations, Holt, Rinehart and Winston, New York, 1961.

39. W. R. Thompson, On the reproduction of organisms with overlapping generations, Bull. Ent. Res., 22 (1931) 147–172.

40. M. B. Usher and M. H. Williamson, A deterministic matrix model for handling the birth, death, and migration processes of spatially distributed populations, Biometrics, 26 (1970) 1–12.

41. P. F. Verhulst, Notice sur la loi que la population suit dans son accroissement, Corresp. math. phys., A. Quelelet, 10 (1838) 113–121.

42. V. J. Vernadsky, The Biosphere, Leningrad (Russian) 1926.

43. V. Volterra, Variation and fluctuations of the number of individuals in animal species living together, Animal Ecology (Chapman, R. N.) McGraw-Hill, New York.

44. S. D. Wicksell, Nuptiality, fertility, and reproductivity, Skand. Aktuarietidskrift (1931) 125–157.

45. G. U. Yule, The growth of population and the factors which control it, J. Roy. Statist. Soc., 88 (1925) 1–58.

THE ISOGONIC AND ERDÖS-MORDELL POINTS OF A SIMPLEX

ROBERT SPIRA, Michigan State University

1. Introduction. We start with a summary of previous work, first taking up the isogonic point and then the Erdös-Mordell inequality.

Fermat proposed the problem: Find the point whose distance from three fixed points has a minimum sum (Johnson [19, p. 221]). The solution is a point, in the interior of the triangle formed by the three given points, except when one angle is $\geqq 120°$, when the solution is at the vertex of this angle. This point is called the isogonic center, and if it is an interior point it subtends equal angles of 120° with all three sides. Steiner [20] gave a discussion of isogonic points. Courant and Robbins [21, Chap. VII] discussed isogonic points.

Given a triangle ABC and a point O within the triangle, let P, Q, R be the points of intersection of perpendiculars from O to the sides. Erdös conjectured

$$OA + OB + OC \geqq 2(OP + OQ + OR) \tag{1}$$

with equality only in the case of an equilateral triangle and O the center. Mordell [2] gave a proof. Barrow [2] gave the stronger result:

$$OA + OB + OC \geqq 2(OE + OF + OG), \tag{2}$$

E, F, and G being the intersections of the interior bisectors of the angles AOB, BOC, and COA with the sides, with the same case of equality as (1).

L. Fejes Tóth [4, pp. 11–14] gave a proof. D. K. Kazarinoff [5] gave a proof, and remarked that with a convention as to signs, the inequality holds for P in the circumcircle. (Without a sign convention, the inequality also holds for some set outside the triangle.) N. D. Kazarinoff [6] showed that for any tetrahedron whose circumcenter is not an exterior point,

$$\sum OA / \sum OP > 2\sqrt{2}, \tag{3}$$

and $2\sqrt{2}$ is the greatest lower bound. He showed the same inequality for trirectangular tetrahedra. A case of equality is given by a degenerate tetrahedron, and a similar case of degenerate equality is found for the conjecture for the n-simplex

$$\sum OA / \sum OP > \frac{(n+1)}{2}\sqrt{2}. \tag{4}$$

Eggleston [7], Veldkamp [8], Brabant [9], and Bankoff [10] gave proofs of (1).

Fejes Tóth [3] conjectured, for convex polygons of n sides,

$$\cos\frac{\pi}{n}\sum OA \geqq \sum OP. \tag{5}$$

Florian [11] proved, for $n=4$,

$$\cos\frac{\pi}{n}\sum OA \geqq \sum OE, \tag{6}$$

$$\cos\frac{\pi}{n}\sum OA \geqq \sum (OA\cdot OB)^{1/2}\cos\left(\frac{1}{2}AOB\right), \tag{6'}$$

where (6′) is stronger than (6). Lenhard [12] showed (6) and (6′), and showed equality holds only for regular n-gons and O the center, and also showed that the result holds when all sides are visible from O. This generalizes Barrow's [2] result. N. D. Kazarinoff [13] gave an exposition of the proof of D. K. Kazarinoff [5].

Oppenheim [14] showed (1) by showing inequalities homogeneous in the OA's and OP's are preserved under certain geometrical transformations, and by showing one of the transformed inequalities. Oppenheim [15] applied a similar technique to (2), and generalized (2) to

$$\lambda OA + \mu OB + \nu OC \geqq 2\sum \frac{\mu\nu(OB+OC)}{\mu OB+\nu OC}OE, \tag{7}$$

where λ, μ, and ν are positive, with equality if and only if $\lambda OA=\mu OB=\nu OC$ and the angles BOC, COA, and AOB are 120°. Mordell [16] gave proofs of (1) and (6′) for $n=3$.

Steinig [17] showed that in a triangle the points O such that $\sum OP=$constant form a line perpendicular to the line joining the incenter and circumcenter. Next, dividing the triangle into two parts by such a perpendicular through the incenter, he showed that in the part containing the circumcenter,

$$\sum OA \geqq 2\sum OP \geqq 6r, \tag{8}$$

and on the other part,

$$\sum OA \geqq 6r \geqq 2\sum OP, \tag{9}$$

where r is the inradius.

Oppenheim [18] extended (1), (2), and (7) to spherical triangles.

2. Results for a Simplex. Let the vertices of a non-degenerate n-simplex S in real n-space be $A_1, \ldots, A_{n+1}$, with $A_i=(a_{i1}, a_{i2}, \ldots, a_{in})$. Let $\overline{S}$ be the closure of S. Let $X=(x_1, x_2, \ldots, x_n)=(x_i)$ be a point in real n-space and define $R(X)=\sum_{j=1}^{n+1}|X-A_i|$. Let B_i be the hyperplane containing all the vertices of S except A_i. Let $r_i(X)=0$ denote the normalized equation of this hyperplane B_i (Schreier and Sperner [1, p. 142]), so that for points inside $\overline{S}$ we have

$$r_i(X) = b_i + \sum_{j=1}^{n} b_{ij}x_j = \text{distance of } X \text{ to } B_i, \tag{10}$$

and

$$\sum_{j=1}^{n} b_{ij}^2 = 1, \tag{11}$$

with $r_i(A_i)>0$. Define

$$r(P) = \sum_{j=1}^{n+1} r_j(P), \tag{12}$$

so $r(P)$ in and on the boundary of S is the sum of the distances to the faces. Hence, for some $c, c_1, \ldots, c_n$,

$$r(P) = c + \sum_{j=1}^{n} c_j x_j, \qquad (\text{in } \bar{S}). \tag{13}$$

Finally, we set $EM(P)=R(P)/r(P)$, the Erdös-Mordell ratio.

LEMMA. *Let $C_1, C_2, \ldots, C_k$ be points of n-space not all lying on the line $x(1,0,\ldots,0)(=xe_1)$. Then the function*

$$f(x) = |xe_1 - C_1| + |xe_1 - C_2| + \cdots + |xe_1 - C_k|$$

is not a linear function of x in any interval.

Proof. If $f(x)=Ax+B$, on some interval, then, setting $C_j=(c_{ij})$,

$$A = f'(x) = \sum_{j=1}^{k} (x - c_{ij})/|xe_1 - C_j|.$$

Now, putting $u=x-c_{1j}$ and $a^2=c_{2j}^2+c_{3j}^2+\cdots+c_{nj}^2$, for $a\neq 0$,

$$\frac{u}{(u^2+a^2)^{1/2}} = \begin{cases} -\dfrac{1}{(1+a^2/u^2)^{1/2}}, & u<0 \\ 0, & u=0 \\ \dfrac{1}{(1+a^2/u^2)^{1/2}}, & u>0, \end{cases} \tag{14}$$

and hence is monotone increasing for $-\infty<x<\infty$. If $a=0$, the function is -1 for $u<0$, and $+1$ for $u>0$. Without loss of generality, we can take the interval not to contain any of the points C_j nor the origin. Hence $f'(x)$ is a sum of constant functions and strictly increasing functions, and contains at least one strictly increasing function. But this contradicts $f'(x)=A$, so the lemma is proved.

PROPOSITION 1. *For $n\geqq 2$, the function $R(P)$ is not a linear function on any interval of any line L.*

Proof. Since $n\geqq 2$, there are at least three A_j's, and by the hypothesis of nondegeneracy, these three A_j's are not collinear. Next, by a rigid motion, (Schreier and Sperner [1, pp. 148–157]) move the points and the line L so that the line L coincides with xe_1. Since distances are preserved by a rigid motion, we can apply the lemma.

PROPOSITION 2. *The set of P's such that* $R(P)\leqq k$ *is convex and compact.*

Proof. If $R(P)\leqq k$ and $R(Q)\leqq k$, we have for $0\leqq\lambda\leqq 1$,

$$\begin{aligned} R(\lambda P+(1-\lambda)Q) &= \sum_{j=1}^{n+1} |\lambda P+(1-\lambda)Q-A_j| \\ &\leqq \sum_{j=1}^{n+1} |\lambda P-\lambda A_j| + |(1-\lambda)Q-(1-\lambda)A_j| \qquad (15)\\ &= \lambda R(P)+(1-\lambda)R(Q) \leqq \lambda k+(1-\lambda)k = k, \end{aligned}$$

so the set is convex. Since $R(P)$ is a continuous function of P, it is easy to see that $\{P|R(P)\leqq k\}$ is closed. For the boundedness, we have

$$R(P) \geqq |P-A_1| \geqq |P|-|A_1|, \tag{16}$$

so for P in the set, $|P|\leqq|A_1|+R(P)\leqq|A_1|+k$.

PROPOSITION 3. *For* $n\geqq 2$, *the function* $R(P)$ *assumes a global minimum at a single point which lies in* $\bar{S}$. *There are no other local minima.*

Proof. Let P_0 be a fixed point and let $k_0=R(P_0)$. Outside the compact set $\{P|R(P)\leqq k_0\}$, $R(P)>k_0$, so the minimum of the continuous function $R(P)$ on the compact set is a global minimum. Let D be the intersection of all nonempty sets $\{P|R(P)\leqq k\}$ with $k\leqq k_0$. The set D is clearly nonempty, compact, and convex, and for some k_D, $R(P)=k_D$ for all points in D, and k_D is the global minimum for $R(P)$. If D contained two points, it would contain the line segment between, and $R(P)$ would be constant on a segment, which violates Proposition 1. Thus, there exists exactly one global minimum point, say $P_{\min}$.

If $P_{\min}$ were outside $\bar{S}$, then for say hyperplane B_1, $P_{\min}$ and A_1 would be on opposite sides. By a rigid motion, take B_1 into the plane $x_1=0$, where the image A_1^* of A_1 has positive first coordinate and the image $P^*_{\min}$ of $P_{\min}$ has negative first coordinate. Let $P_2=P^*_{\min}$ except for the sign of its first coordinate. Then

$$\begin{aligned} R(P_2) &= |P_2-A_1|+|P_2-A_2|+\dots+|P_2-A_{n+1}| \\ &= R(P_{\min})+|P_2-A_1|-|P_{\min}-A_1| \\ &< R(P_{\min}), \end{aligned}$$

since, setting $P_{\min}=(p_1,\cdots,p_n)$,

$$\begin{aligned} |P_{\min}-A_1|^2 &= (p_1-a_{11})^2+(p_2-a_{12})^2+\dots+(p_n-a_{1n})^2 \\ &> (-p_1-a_{11})^2+(p_2-a_{12})^2+\dots+(p_n-a_{1n})^2 = |P_2-A_1|^2. \end{aligned}$$

But $P_{\min}$ is the global minimum, so indeed $P_{\min}\in\bar{S}$.

Finally, if P_1 were a second local minimum point, with $R(P_1)=k_1>k_D$, then the segment $P_1P_{\min}$ would be in the set $\{P|R(P)\leqq k_1\}$. But in an interval along this segment, by the local minimum property, $R(P)\geqq R(P_1)=k_1$. Thus, $R(P)$ would be constant along a segment, contradicting Proposition 1. Hence there cannot be a second local minimum.

In Section 3, we show for $n=3$ that the minimum is either at an interior point or a vertex. For higher dimensions, the scientific situation could be explored by calculations.

Next, we take up the Erdös-Mordell point.

PROPOSITION 4. *The set of P's in* $\bar{S}$ *satisfying* $EM(P)\leqq k$ *is convex and compact.*

Proof. In $\bar{S}$ we can use the expression (13) for $r(P)$. Let $EM(P)\leqq k$ and $EM(Q)\leqq k$. Then, setting $P=(p_j)$, $Q=(q_j)$,

$$\begin{aligned} EM(\lambda P + (1-\lambda)Q) &= \frac{R(\lambda P + (1-\lambda)Q)}{r(\lambda P + (1-\lambda)Q)} \\ &= \frac{R(\lambda P + (1-\lambda)Q)}{\lambda r(P) + (1-\lambda)r(Q)} \qquad (17) \\ &\leqq \frac{\lambda R(P) + (1-\lambda)R(Q)}{\lambda r(P) + (1-\lambda)r(Q)}, \end{aligned}$$

using (15).

Next, since $R(P)\leqq kr(P)$ and $R(Q)\leqq kr(Q)$, we have

$$\lambda R(P) + (1-\lambda)R(Q) \leqq k(\lambda r(P) + (1-\lambda)r(Q)). \qquad (18)$$

Thus, the ratio ending the string of equalities and inequalities of (17) is $\leqq k$, and convexity is shown. As in the proof of Proposition 2, $EM(P)$ is continuous, so the set in question is closed, and clearly bounded as it is included in $\bar{S}$, so compactness is shown.

PROPOSITION 5. *For* $n\geqq 2$, *the function* $EM(P)$ *assumes its minimum on* $\bar{S}$ *at exactly one point. There are no other local minima in* $\bar{S}$.

Proof. Let P_0 be a point in $\bar{S}$, and let $k_0=EM(P_0)$. Let D be the intersection of all nonempty sets $\bar{S}\cap\{P\mid EM(P)\leqq k\}$ with $k\leqq k_0$. The set D is clearly nonempty, compact, and convex, and for some k_D, $EM(P)=k_D$ for all $P\in D$, and k_D is the minimum of $EM(P)$ over $\bar{S}$. If D contained two points, $EM(P)$ would be constant on a segment. By a rigid motion bring this segment into coincidence with the x_1 axis. Then $EM(xe_1)=k_D$, or $R(xe_1)=k_Dr(xe_1)=k_D(c+c_1x)$, and $R(P)$ would be a linear function on an interval, which contradicts Proposition 1. Thus, D contains only a single point. The argument that there are no other local minima is similar to that given for Proposition 3.

V. P. Sreedharan remarked that another proof may be obtained from the theorem: If f and g are positive on a convex set E, and if f is strictly convex (convex) and g is concave (strictly concave) then the ratio f/g assumes at most one maximum.

In the case of a triangle, the Erdös-Mordell point can lie on the boundary. For an isosceles right triangle, it is an interior point, but for other right triangles

it can lie on the hypotenuse. For obtuse isosceles triangles, setting γ = one half the obtuse angle, the point will be a boundary point if and only if

$$4 \sin^2 \gamma \geqq 2 \sin \gamma + \cos \gamma.$$

Let α and β be two angles of a triangle; it would be interesting to find the region in the $\alpha \cdot \beta$ plane, where the Erdös-Mordell point is an interior point. A similar question could also be asked for $n=3, 4, \cdots,$. One would also like to know whether the Erdös-Mordell point could lie at a vertex.

3. Results for a Tetrahedron. We now sharpen our results for the case $n=3$.

PROPOSITION 6. *For $n=3$, the function $R(P)$ assumes its minimum at either a vertex or an interior point.*

Proof. First let P lie on the interior of side $A_1A_2A_3$ and let A_4P make an angle φ with the normal to $A_1A_2A_3$ at P, $0<\varphi\leqq\pi/2$. Let P_ϵ be a distance ϵ along the normal into the interior of the tetrahedron. Let $|P-A_1|=a$, $|P-A_2|=b$, $|P-A_3|=c$, $|P-A_4|=d$. We have then

$$\begin{aligned} R(P_\epsilon) &= (d^2+\epsilon^2-2\epsilon d\cos\varphi)^{1/2}+(a^2+\epsilon^2)^{1/2}+(b^2+\epsilon^2)^{1/2}+(c^2+\epsilon^2)^{1/2} \\ &= (d-\epsilon)(1+(1-\cos\varphi)2d\epsilon/(d-\epsilon)^2)^{1/2} \\ &\qquad + a(1+\epsilon^2/a^2)^{1/2}+b(1+\epsilon^2/b^2)^{1/2}+c(1+\epsilon^2/c^2)^{1/2}. \end{aligned}$$

Next, for $f(x)=(1+x)^{1/2}$, we have by the law of the mean,

$$(1+h)^{1/2} = 1 + h/2(1+\xi)^{1/2} \leqq 1 + h/2$$

for $h>0$. Thus,

$$R(P_\epsilon) \leqq a+b+c+d-\epsilon+\epsilon\frac{d}{d-\epsilon}(1-\cos\varphi)+\frac{\epsilon^2}{2}\left(\frac{1}{a}+\frac{1}{b}+\frac{1}{c}\right).$$

Choose ρ so $1-\cos\varphi<\rho<1$. Choose ϵ so

$$\frac{d}{d-\epsilon}<\frac{\rho}{1-\cos\varphi}$$

for ϵ sufficiently small. Thus,

$$R(P_\epsilon) \leqq R(P)+\epsilon\left\{(\rho-1)+\frac{\epsilon}{2}\left(\frac{1}{a}+\frac{1}{b}+\frac{1}{c}\right)\right\},$$

so if we further restrict ϵ so that

$$\epsilon < 2(1-\rho)\Big/\left(\frac{1}{a}+\frac{1}{b}+\frac{1}{c}\right),$$

then $R(P_\epsilon)<R(P)$, so P cannot be a minimum point. Next, let P lie on an edge, but not at a vertex. Let say P lie on edge A_2A_4 and let the distances to the vertices be a, b, c, and d as above. The angle A_1PA_3 is clearly less than π towards the

interior of the tetrahedron, and we construct a point P_ϵ a distance ϵ along the interior bisector of the angle towards the interior. Let $\angle P_\epsilon PA_1 = \angle P_\epsilon PA_3 = \theta$, $0 < \theta < \pi/2$. Let the acute angle between the segments PP_ϵ and A_2A_4 be ϕ, and suppose without loss of generality $\angle P_\epsilon PA_2 = \phi$. We then have

$$
\begin{aligned}
R(P_\epsilon) &= (a^2 + \epsilon^2 - 2a\epsilon \cos \theta)^{1/2} + (c^2 + \epsilon^2 - 2c\epsilon \cos \theta)^{1/2} \\
&\quad + (b^2 + \epsilon^2 - 2b\epsilon \cos \phi)^{1/2} + (d^2 + \epsilon^2 + 2d\epsilon \cos \phi)^{1/2} \\
&\leqq a - \epsilon + \epsilon \frac{a}{a - \epsilon}(1 - \cos \theta) + c - \epsilon + \epsilon \frac{c}{c - \epsilon}(1 - \cos \theta) \\
&\quad + b - \epsilon + \epsilon \frac{b}{b - \epsilon}(1 - \cos \phi) + d + \epsilon - \epsilon \frac{d}{d + \epsilon}(1 - \cos \phi) \\
&\leqq R(P) - \epsilon \left\{ 2 - (1 - \cos \theta)\left(\frac{a}{a - \epsilon} + \frac{b}{b - \epsilon} \right) \right. \\
&\qquad \left. + (1 - \cos \phi)\left(\frac{d}{d + \epsilon} - \frac{b}{b - \epsilon} \right) \right\} < R(P)
\end{aligned}
$$

since

$$
\frac{d}{d + \epsilon} - \frac{b}{b - \epsilon} \to 0 \quad \text{and} \quad \frac{a}{a - \epsilon} + \frac{b}{b - \epsilon} \to 2
$$

and $1 - \cos \theta < 1$. Thus, Proposition 6 is proved.

If the minimum is at a vertex, it occurs at the vertex whose incident edge lengths have the least sum. If two or more such vertices exist, the minimum is of course in the interior.

It is easy to see that in a triangle the angle subtended at the sides is 120°. Indeed the minimum point can be thought of as lying on an ellipse with two vertices as foci, so the line to the third vertex is perpendicular to the tangent to the ellipse. Since the lines to the foci make equal angles with the tangent, two of the subtended angles, by addition, are equal. Rotating the argument, we get all three angles equal. Hence also, if one of the angles of the triangle is $\geqq 120°$, the minimum occurs at a vertex, and clearly by comparison of the lengths of the sides, at the obtuse vertex.

For the tetrahedron, it is easy to show that a vertex for which there are two acute face angles cannot be a minimum point. One simply moves a distance ϵ down the edge dividing the two faces. If one constructs a line making an angle $\theta < \pi/2$ with each of the three edges as a vertex, it is easy to see that by moving a distance ϵ along this line, the sum of the distances will decrease if $\cos \theta > 1/3$. (Such a line may lie outside the tetrahedron.) The proofs of these statements are similar to the proofs given above, for ϵ small enough.

It may be that $\cos \theta \leqq 1/3$ will be the exact condition for the minimum to lie at a vertex. The present author has not been able to prove this.

An interesting theorem, which could possibly be of use, is the following:

THEOREM. *A point P which has the minimum sum of distances to the vertices of a tetrahedron, and which is also an interior point, subtends equal angles at opposite edges.*

Proof. Using the notation as above, let A_1, A_2, A_3, and A_4 be the vertices. Since the minimum occurs at an interior point, we have

$$\left.\frac{\partial R}{\partial x_1}\right|_P = \left.\frac{\partial R}{\partial x_2}\right|_P = \left.\frac{\partial R}{\partial x_3}\right|_P = 0.$$

Let $P=(p_1, p_2, p_3)$. After squaring and rearranging, we obtain

$$\left(\frac{p_j - a_{1j}}{|P - A_1|} + \frac{p_j - a_{2j}}{|P - A_2|}\right)^2 = \left(\frac{p_j - a_{3j}}{|P - A_3|} + \frac{p_j - a_{4j}}{|P - A_4|}\right)^2, \qquad j = 1, 2, 3,$$

and adding these, we obtain

$$1 + 2\cos\theta_{12} + 1 = 1 + 2\cos\theta_{34} + 1,$$

where θ_{ij} is the angle between A_iP and A_jP. Thus $\theta_{12}=\theta_{34}$, and similarly $\theta_{13}=\theta_{24}$ and $\theta_{14}=\theta_{23}$. B. Stewart kindly supplied this proof.

Calculations would probably reveal the situation for higher dimensions. A simple physical model of the situation for three space may be constructed from four fixed rings, numbered from 1 to 4, a movable ring numbered 0 and a piece of thread. Pass the thread successively through 4, 0, 1, 0, 2, 0, 3, 0, 4, and tighten the thread by pulling at the two ends through 4, perturbing slightly ring 0. When stability is reached, the angles may be measured with a protractor.

Using this model, Leroy Kelly supplied a physics proof. The forces on the central ring are in equilibrium and the tensions in the strings are equal. Letting the force vectors be PA_i, we have

$$2|PA_1|\cos\frac{\theta_{12}}{2} = |PA_1 + PA_2| = |PA_3 + PA_4| = 2|PA_3|\cos\frac{\theta_{34}}{2},$$

so $\theta_{12}=\theta_{34}$.

Finally, the author remarks that it is not clear exactly how Barrow's result (2) should be generalized to higher dimensions. For $n=3$, one could generalize to lines which make equal angles with the triples of lines OA_i, to lines which make equal angles with the triples of sides OA_iA_j, or to the diagonals of rhomboids formed with triples of lines OA_i. Algebraically, the rhomboid diagonals appear closest to Barrow's proof, but the constant for the degenerate case of Kazarinoff [6] is smaller, $4/5^{1/2}$ rather than $8^{1/2}$. In this degenerate case, the lines making equal angles with the triples of sides give a constant of 2, while the lines making equal angles with the triples of edges give the Kazarinoff constant of $8^{1/2}$.

References

1. O. Schreier and E. Sperner, Einführung in die analytische Geometrie und Algebra, Teubner, Berlin, 1936.
2. Advanced Problem 3740, this MONTHLY, 44 (1937) 252–4.
3. L. Fejes Tóth, Inequalities concerning polygons and polyhedra, Duke Math. J., 15 (1948) 817–822.
4. L. Fejes Tóth, Lagerungen in der Ebene auf der Kugel und im Raum, Springer, Berlin, 1953.
5. D. K. Kazarinoff, A simple proof of the Erdös-Mordell inequality for triangles, Michigan Math. J., 4 (1957) 97–98.
6. N. D. Kazarinoff, D. K. Kazarinoff's inequality for tetrahedra, Michigan Math. J., 4 (1957) 99–104.
7. H. G. Eggleston, A triangle inequality, Math. Gaz., 42 (1958) 54–55.
8. G. R. Veldkamp, The Erdös-Mordell inequality, Nieuw Tijdschr. Wisk., 45 (1957/58) 193–196.
9. H. Brabant, The inequality of Erdös-Mordell again, Nieuw Tijdschr. Wisk., 46 (1958/59), 87.
10. L. Bankoff, An elementary proof of the Erdös-Mordell Theorem, this MONTHLY, 65 (1958) 521.
11. A. Florian, Zu einem Satz von P. Erdös, Elem. Math., 13 (1958) 55–58.
12. H. Lenhard, Verallgemeinerung und Verschärfung der Erdös-Mordellschen Ungleichung für Polygone, Arch. Math., 12 (1961) 311–314.
13. N. D. Kazarinoff, Geometric Inequalities, Random House, New York, 1961.
14. A. Oppenheim, The Erdös inequality and other inequalities for a triangle, this MONTHLY, 68 (1961) 226–230.
15. ———, New inequalities for a triangle and an internal point, Ann. Mat. Pura Appl., (4) 53 (1961) 157–164.
16. L. J. Mordell, On geometric problems of Erdös and Oppenheim, Math. Gaz., 46 (1962) 213–215.
17. J. Steinig, À comparison of two inequalities for the triangle, Acta. Math. Acad. Sci. Hungar., 16 (1965) 19–22.
18. A. Oppenheim, Some inequalities for a spherical triangle and an internal point, Publications de la Faculté d'électrotechnique de l'université à Belgrade, Série: Mathématiques et Physique, No. 203 (1967) 13–16.
19. R. A. Johnson, Advanced Euclidean Geometry, Dover, New York, 1960.
20. J. Steiner, Über den Punkt der kleinsten Entfernung, Monatsb. Akad. Wiss. Berlin, (1837) 144. Werke, II, pp. 93–95, pp. 729–731.
21. R. Courant and H. Robbins, What is Mathematics?, Oxford Univ. Press, New York, 1941.

PARTIAL SUMS OF THE HARMONIC SERIES

R. P. BOAS, JR., Northwestern University, and J. W. WRENCH, JR., Naval Ship Research and Development Center

1. Introduction. Our problem is to compute the smallest integer $n=n_A$ for which the sum S_n of the first n terms of the harmonic series $\sum 1/k$ exceeds a given number A, say $A=100$. (Clearly, when A is this large there is no possibility of adding up the individual terms.) The problem has been mentioned in a number of places. Hardy ([5], p. 69) claims to give n_A for $A=3$, 5, 10, and approxi-

mations for 10^2, 10^3, 10^6; but except for $A=3$ his values are too small, as we shall see. Gardner [4] quotes a calculation by D. Asimov as showing that $2^{143}<n_{100}<2^{144}$. Comtet [1] gives a formula from which n_A can be calculated with an error of at most 1, but gives no numerical results.

We shall establish a theorem that lets us calculate n_A exactly for each integer A that we have tested, and in particular for $A=100$. It is "almost" true that $n_A=[e^{A-\gamma}+\frac{1}{2}]$, where the brackets denote integral part and γ is Euler's constant; it is in fact true for all the integers A that we have examined, although it is false for some nonintegral values of A. We note incidentally that S_n is never an integer for $n>1$ (see [7]).

We append two numerical tables. The first consists of decimal approximations to $e^{A-\gamma}$, rounded to 16 significant figures, for $A=1(1)20$, supplemented by the values of $e^{-\gamma}$ to 50 figures and $e^{100-\gamma}$ to 49 figures. The second gives to 10 decimal places the partial sums S_n of the harmonic series for $n=n_A-1$ and $n=n_A$, respectively, for $A=1(1)20$.

The entries in the first table were initially computed to 20 significant figures by multiplying the appropriate entries in the fundamental tables of Van Orstrand [9] by a 20-place approximation to $e^{-\gamma}$ ([3], p. 137). The tabulated value of $e^{100-\gamma}$ is an abridgment of an approximation calculated to about 116 figures from a comparable approximation to e^{100} published by Uhler [8] and an unpublished value of $e^{-\gamma}$ computed to about 170 decimals by Wrench.

Entries in the second table corresponding to values of n exceeding 100 were calculated by the Euler-Maclaurin asymptotic formula; the entries for smaller values of n were calculated either by direct summation or by using tables of the psi function [2].

We shall prove the following theorem:

THEOREM 1. *Let n_A denote the smallest integer n for which $\sum_{k=1}^{n} 1/k$ exceeds $A(\geqq 3)$. Let $e^{A-\gamma}=m+\delta$, where m is an integer and $0<\delta<1$. Then $n_A=m$ if $\delta<\frac{1}{2}-(10\,n)^{-1}$, and $n_A=m+1$ if $\delta>\frac{1}{2}+n^{-1}$.*

In other words, $n_A=[e^{A-\gamma}+\frac{1}{2}]$ unless $e^{A-\gamma}+\frac{1}{2}$ is too close to an integer.

The numerical coefficients of n^{-1} could be improved by more attention to detail.

We can state Theorem 1 in another way. Let us divide the positive real numbers A into classes C_δ according to the value of δ, the fractional part of $e^{A-\gamma}$. Then, for a given δ, all sufficiently large elements of C_δ have $n_A=[e^{A-\gamma}+\frac{1}{2}]$. We state this more precisely as follows:

THEOREM 2. *If m and δ are the integral and fractional parts of $e^{A-\gamma}$, then $n_A=[e^{A-\gamma}+\frac{1}{2}]$ provided that $m>2/(1-2\delta)$.*

2. Qualitative results. We first present a rather crude analysis based on the simplest version of the Euler-Maclaurin formula ([6], p. 523). Although this is not sufficiently precise to let us establish Theorem 1, it has enabled us to determine n_A for all the integers A that we have examined. The proofs of Theorems

1 and 2 (given in section 3) require an additional term in the formula; greater precision could be had at the expense of taking still more terms.

The same methods apply to any series for which the analogue of the Euler constant (or the sum of the series, if it converges) is known with sufficient accuracy.

According to the Euler-Maclaurin formula, the partial sums S_n satisfy

$$S_n - \log n = \tfrac{1}{2}(n^{-1} + 1) - \int_1^n P_1(t)t^{-2}dt,$$

where $P_1(t) = t - [t] - \frac{1}{2}$. As $n \to \infty$, the left-hand side approaches γ, and thus

$$\gamma = \tfrac{1}{2} - \int_1^\infty P_1(t)t^{-2}dt = \tfrac{1}{2} - \int_1^n P_1(t)t^{-2}dt - \int_n^\infty P_1(t)t^{-2}dt,$$

so that

$$S_n = \log n + (2n)^{-1} + \gamma + \int_n^\infty P_1(t)t^{-2}dt.$$

By the second mean-value theorem ([**10**], p. 163),

$$R_n \equiv \int_n^\infty P_1(t)t^{-2}dt = n^{-2}\int_n^c P_1(t)dt, \qquad c \geqq n.$$

Since P_1 has period 1, mean-value zero, and maximum absolute value $\frac{1}{2}$, we have $0 \geqq R_n \geqq \frac{1}{8}n^{-2}$. Consequently $S_n > A$ if

$$\log n > A - \gamma - (2n)^{-1} + (8n^2)^{-1}, \tag{1}$$

and $S_n < A$ if

$$\log n < A - \gamma - (2n)^{-1}. \tag{2}$$

Now let $m = [e^{A-\gamma}]$, that is, $e^{A-\gamma} = m + \delta$, where m is an integer and $0 < \delta < 1$. Then $S_n > A$ if

$$\log n > \log(m + \delta) - (2n)^{-1} + (8n^2)^{-1}, \tag{3}$$

and $S_n < A$ if

$$\log n < \log(m + \delta) - (2n)^{-1}. \tag{4}$$

In a specific numerical case it is usually easy to check (3) and (4) for $n = m$, $m \pm 1$.

We now establish a qualitative version of Theorem 2. There are four things to prove: (i) $S_{m-1} < A$ when A is sufficiently large; (ii) if $0 < \delta < \frac{1}{2}$, then $S_m > A$ when A is sufficiently large; (iii) if $\frac{1}{2} < \delta < 1$, then $S_m < A$ when A is sufficiently large; (iv) if $\frac{1}{2} < \delta < 1$, then $S_{m+1} > A$ when A is sufficiently large. We observe, in fact, that the appropriate pair of statements hold for all the values of A in Tables I and II.

(i) It is enough, by (4), to show that

$$\log(m-1) < \log(m+\delta) - (2m-2)^{-1}.$$

If we expand the logarithms in series, we see that this means

$$0 < \frac{\delta}{m} - \frac{\delta^2}{2m^2} + \cdots + \frac{1}{m} + \frac{1}{2m^2} + \cdots - \frac{2}{2(m-1)}\,.$$

The right-hand side is certainly positive if $m^{-1}-\frac{1}{2}(m-1)^{-1}>0$, which is true if $m>2$ and so if $A>2$; for this part the size of δ is irrelevant.

(ii) If $0<\delta<\frac{1}{2}$, we shall verify (3) for $n=m$, that is,

$$\log m > \log(m+\delta) - (2m)^{-1} + (8m^2)^{-1}.$$

This inequality can be written

$$\log m > \log m + \left\{\frac{\delta}{m} - \frac{\delta^2}{2m^2} + \cdots\right\} - \frac{1}{2m} + \frac{1}{8m^2},$$

i.e., $0>(\delta-\frac{1}{2})m^{-1}+O(m^{-2})$. Thus, if δ is fixed, $0<\delta<\frac{1}{2}$, and A (hence m) is large enough, we have $S_m>A$.

(iii) In the same way, (4) holds for $n=m$ if

$$0 < (\delta - \tfrac{1}{2})m^{-1} + O(m^{-2});$$

and so if $\frac{1}{2}<\delta<1$ and A is large enough, we have $S_m<A$.

(iv) If $\frac{1}{2}<\delta<1$ we verify (3) for $n=m+1$. We need

$$\log(m+1) + \tfrac{1}{2}(m+1)^{-1} > \log(m+\delta) + \tfrac{1}{8}(m+1)^{-2}.$$

Since $\log x+\frac{1}{2}x^{-1}$ increases, this inequality follows from

$$\tfrac{1}{2}(m+\delta)^{-1} > \tfrac{1}{8}(m+1)^{-2},$$

which is true for large enough m provided that $\delta<1$.

To see that the first n for which $S_n>A$ is not necessarily $m+1$ when $\delta>\frac{1}{2}$, we may easily calculate that if we select m and then (perversely) choose to consider an A such that $e^{A-\gamma}=m+\frac{1}{2}+\epsilon/m$ with $0<\epsilon<\frac{1}{8}$, then $S_n>A$ for $n=m$, but not for $n=m-1$. In other words, the "almost" theorem of Section 1 fails for such an A.

3. Proof of Theorem 1. To establish Theorem 1 (and hence Theorem 2), we have to take an additional term in the Euler-Maclaurin formula ([6], p. 526). We have

$$S_n = \log n + \frac{1}{2}(1+n^{-1}) + \frac{1}{12}(1-n^{-2}) - 6\int_1^n P_3(t)t^{-4}dt,$$

and

$$\gamma = \frac{1}{2} + \frac{1}{12} - 6\int_1^\infty P_3(t)t^{-4}dt,$$

where it is sufficient for our purposes to know that $|P_3(t)| \leqq \frac{1}{2}$. We then have

$$S_n = \log n + (2n)^{-1} + \gamma + R,$$

with $|R| < (12n^2)^{-1} + n^{-3}$. Consequently $S_n > A$ if

$$\log n + (2n)^{-1} > A - \gamma + (12n^2)^{-1} + n^{-3},$$

and $S_n < A$ if

$$\log n + (2n)^{-1} < A - \gamma - (12n^2)^{-1} - n^{-3}.$$

Again let $e^{A-\gamma} = m + \delta$, $0 < \delta < 1$. It is enough to prove:

(i) $S_{m-1} < A$, which has already been done in Section 2;
(ii) if $0 < \delta < \frac{1}{2} - (10n)^{-1}$, then $S_m > A$;
(iii) if $\frac{1}{2} + n^{-1} < \delta < 1$, then $S_m < A$;
(iv) $S_{m+1} > A$.

Proof of (ii). We have $S_m > A$ if $\log m > \log(m+\delta) - (2m)^{-1} + (12m^2)^{-1} + m^{-3}$. That is,

$$0 > \left(\delta - \frac{1}{2}\right)m^{-1} + \left(\frac{1}{12} - \frac{1}{2}\delta^2\right)m^{-2} + \left(\frac{1}{3}\delta^3 + 1\right)m^{-3}$$
$$- \frac{1}{4}\delta^4 m^{-4} + \frac{1}{5}\delta^5 m^{-5} - \cdots.$$

Let $\delta < \frac{1}{2} - \epsilon m^{-1}$. We then have $S_m > A$ if

$$0 > \left(-\epsilon + \frac{1}{12}\right)m^{-2} + \frac{25}{24}m^{-3} - \frac{1}{4}\delta^4 m^{-4} + \cdots. \tag{5}$$

The sum of the positive terms on the right is less than

$$\frac{25}{24}m^{-3} + \frac{1}{5}\left\{(2m)^{-5} + (2m)^{-7} + \cdots\right\} < m^{-3}\left(\frac{25}{24} + \frac{1}{100}\right),$$

and so (5) certainly holds if $m > 2$ and $\epsilon \geqq 1/10$.

Proof of (iii). We have $S_m < A$ if $\log m < \log(m+\delta) - (2m)^{-1} - (12m^2)^{-1} - m^{-3}$, and hence if

$$0 < \left(\delta - \frac{1}{2}\right)m^{-1} - \left(\frac{1}{2}\delta^2 + \frac{1}{12}\right)m^{-2} - \left(1 - \frac{1}{3}\delta^3\right)m^{-3} - \frac{1}{4}\delta^4 m^{-4} + \cdots.$$

If $\frac{1}{2} + m^{-1} < \delta < 1$, the first term on the right is at least m^{-2} and the rest is at least

$$-m^{-2}\left\{\left(\frac{1}{2}\delta^2 + \frac{1}{12}\right) + \left(1 - \frac{1}{3}\delta^3\right)m^{-1} + \frac{1}{4}\delta^4 m^{-2} + \cdots\right\}$$

$$\geqq -m^{-2}\left\{\frac{7}{12}+m^{-1}(1+m^{-1}+m^{-2}+\cdots)\right\}$$

$$=-m^{-2}\left\{\frac{7}{12}+\frac{1}{m-1}\right\}>-m^{2}\qquad \text{if } m>3.$$

Proof of (iv). We have $S_{m+1}>A$ if

$$\log(m+1)+\frac{1}{2}(m+1)^{-1}>A-\gamma+\frac{1}{12}(m+1)^{-2}+(m+1)^{-3}. \tag{6}$$

Since $\log x+(2x)^{-1}$ increases, the left side of (6) exceeds $A-\gamma+\frac{1}{2}(m+\delta)^{-1}$, which exceeds the right side of (6) if $m>3$.

TABLE 1. Table of $e^{A-\gamma}$ to 16 Significant Figures.
For each $A>0$ in this table, $n_A=[e^{A-\gamma}+\frac{1}{2}]$.

A			$e^{A-\gamma}$		
0	0.56145	94835	66885	16982	41432
	14790	88078	67657	10386	92515
1		1.	52620	51115	95864
2		4.	14865	56213	52346
3		11.	27721	51880	5655
4		30.	65464	91213	1648
5		83.	32797	56642	6262
6		226.	50892	20504	426
7		615.	71508	67935	645
8		1673.	68713	19390	30
9		4549.	55331	72756	03
10		12366.	96810	99558	4
11		33616.	90468	64254	6
12		91380.	22113	81500	6
13	2	48397.	19460	04024	
14	6	75213.	58032	24792	
15	18	35420.	80571	9367	
16	49	89191.	02376	2615	
17	135	62027.	29860	490	
18	368	65412.	36286	321	
19	1002	10580.	52462	05	
20	2724	00600.	05940	78	
100	1509	26886	22113	78832	36935
	63264	53810	14498	59497.	36410

TABLE 2. Selected partial sums of the harmonic series

n	S_n	n	S_n
1	1.00000 00000	33616	10.99998 79618
2	1.50000 00000	33617	11.00001 77086
3	1.83333 33333	91379	11.99999 21084
4	2.08333 33333	91380	12.00000 30517
10	2.92896 82540	248396	12.99999 72037
11	3.01987 73449	248397	13.00000 12295
30	3.99498 71309	675213	13.99999 98810
31	4.02724 51954	675214	14.00000 13621
82	4.99002 00799	1835420	14.99999 98334
83	5.00206 82727	1835421	15.00000 03783
226	5.99996 14220	4989190	15.99999 98950
227	6.00436 67083	4989191	16.00000 00955
615	6.99965 07205	13562026	16.99999 99411
616	7.00127 40971	13562027	17.00000 00148
1673	7.99988 82004	36865411	17.99999 99766
1674	8.00048 55720	36865412	18.00000 00037
4549	8.99998 82827	100210580	18.99999 99998
4550	9.00020 80629	100210581	19.00000 00097
12366	9.99996 21479	272400599	19.99999 99979
12367	10.00004 30083	272400600	20.00000 00016

References

1. L. Comtet, Problem 5346, this MONTHLY, 74 (1967) 209.
2. H. T. Davis, Tables of the mathematical functions, vol. I, rev. ed., Principia Press of Trinity University, San Antonio, Texas, 1963.
3. A. Fletcher, J. C. P. Miller, L. Rosenhead, L. J. Comrie, An index of mathematical tables, v. I., 2nd ed., Addison-Wesley, Reading, Mass., 1962.
4. M. Gardner, Mathematical games, Scientific American, January, 1965, p. 116.
5. G. H. Hardy, Orders of infinity, 2nd edition, Cambridge University Press, 1924.
6. K. Knopp, Theory and application of infinite series, London-Glasgow, 1928.
7. G. Pólya and G. Szegö, Aufgaben und Lehrsätze aus der Analysis, Springer Verlag, Berlin, 1925, v. 2, pp. 159, 380–381, problems VIII, 250, 251.
8. H. S. Uhler, A new table of reciprocals of factorials and some derived numbers, Trans. Conn. Acad. Arts and Sci., 32 (1937) 381–434; p. 433.
9. C. E. Van Orstrand, Tables of the exponential function and of the circular sine and cosine to radian argument, Mem. Nat. Acad. Sci., 14(5) (1921).
10. D. V. Widder, Advanced Calculus, 2nd. ed., Prentice Hall, Englewood Cliffs, N. J., 1961.

CORRECTION TO "PATTERNS OF VISIBLE AND NONVISIBLE LATTICE POINTS"

FRITZ HERZOG AND B. M. STEWART, Michigan State University

In the above article (this MONTHLY, 78(1971) 487–496) equation (6′) on page 495 should read as follows:

$$(6') \quad (u_1, u_2, \cdots, u_k) \equiv (-i_1, -i_2, \cdots, -i_k) \bmod Q(i_1, i_2, \cdots, i_k).$$

ON BERNSTEIN'S INEQUALITY AND THE NORM OF HERMITIAN OPERATORS

ANDREW BROWDER, Brown University

Consider the following two propositions:

PROPOSITION A. *If T is a bounded Hermitian operator on an inner product space, then* $\|T^2\|=\|T\|^2$.

PROPOSITION B. *Let f be a trigonometric polynomial of degree N. Then* $|f'(x)| \leqq N \operatorname{Max}\{|f(t)| : t \text{ real}\}$.

Proposition A is quite trivial:

$$\|T\|^2 = \sup\{(Tx, Tx):\|x\| = 1\} = \sup\{(T^2x, x):\|x\| = 1\} \leqq \|T^2\| \leqq \|T\|^2.$$

Proposition B is a well-known inequality of Sergei Bernstein.

At first sight, these propositions appear quite unrelated. However, the natural generalizations of these propositions turn out to be equivalent. We begin by generalizing Bernstein's inequality.

PROPOSITION B′. *Let f be an entire function, and suppose there exist constants M and A such that $|f(\sigma+it)| \leqq Me^{A|t|}$ for each complex number $s=\sigma+it$. Then $|f'(u)| \leqq A \sup\{|f(\sigma)| : \sigma \text{ real}\}$ for every real u.*

It is easy to see that Proposition B′ (also due to Bernstein) contains Proposition B. For if $f(s) = \sum_{-N}^{N} c_n e^{ins}$, clearly $|f(s)| \leqq Me^{N|t|}$, with $M = \sum |c_n|$.

Let us give a proof of B′. Suppose that f is entire, and $|f(s)| \leqq Me^{|t|}$ for all $s=\sigma+it$. Let $F(s) = f(s)/s^2 \cos s$. Then F is meromorphic, with poles at odd multiples of $\frac{1}{2}\pi$ and at 0. If $\zeta=(2k+1)\pi/2$, the residue of F at ζ is $(-1)^{k+1}f(\zeta)/\zeta^2$; the residue of F at 0 is $f'(0)$. Let Γ_n be the square contour with corners at $\pi n(\pm 1 \pm i)$. Then

$$\frac{1}{2\pi i}\int_{\Gamma_n} F(\zeta)d\zeta = f'(0) - \sum_{-n}^{n-1} \frac{4}{\pi^2}(-1)^k \frac{f((2k+1)\pi/2)}{(2k+1)^2}$$

by the residue theorem. Now it is easy to see that $|\cos s| > \frac{1}{3}e^{|t|}$ on Γ_n, so $|F(s)| \leqq 3|M||s|^{-2}$ on Γ_n. It follows that $\int_{\Gamma_n} F(\zeta)d\zeta \to 0$ as $n \to \infty$, and hence that

$$f'(0) = \frac{4}{\pi^2}\sum_{-\infty}^{\infty}(-1)^k \frac{f((2k+1)\pi/2)}{(2k+1)^2},$$

whenever f is entire and $|f(s)|e^{-|t|}$ is bounded. Taking $f(s) = \sin s$, we find that

$$\sum_{-\infty}^{\infty}\frac{1}{(2k+1)^2} = \frac{\pi^2}{4}.$$

In general, we have

$$|f'(0)| \leqq \frac{4}{\pi^2}\sum_{-\infty}^{\infty}\frac{|f((2k+1)\pi/2)|}{(2k+1)^2} \leqq \sup\{|f(\sigma)| : \sigma \text{ real}\}\cdot\frac{4}{\pi^2}\sum_{-\infty}^{\infty}\frac{1}{(2k+1)^2}$$

$$= \sup\{|f(\sigma)| : \sigma \text{ real}\}.$$

Now if $|f(s)|e^{-A|t|}$ is bounded, then $|f(s/A)|e^{-|t|}$ is bounded, so $|f'(0)| \leqq A \sup\{|f(\sigma)| : \sigma \text{ real}\}$. By considering $f(s+u)$, we obtain $|f'(u)| \leqq A \sup\{|f(\sigma)| : \sigma \text{ real}\}$.

How do we generalize Proposition A? The notion of Hermitian operators on normed linear spaces was first introduced by Vidav [1], an equivalent definition given by Lumer [2]. Both definitions are equivalent to the following:

DEFINITION. A bounded operator T on a normed linear space is called **Hermitian** if $\|e^{iuT}\| = 1$ for all real u.

The natural generalization of A is then the following:

PROPOSITION A′. *If T is a bounded Hermitian operator on a normed linear space X, then $\|T^n\| = \|T\|^n$, for every positive integer n.*

Proof. Fix the positive integer n. Let $A = \|T^n\|^{1/n}$. For any positive integer m, we may write $m = nk + l$, where k and l are positive integers and $0 \leqq l < n$. Then $\|T^m\| \leqq \|T^{nk}\| \, \|T^l\| \leqq \|T^n\|^k \|T^l\| = A^{nk}\|T^l\|$, so (if $A > 0$) we have $\|T^m\| \leqq MA^m$, where $M = \operatorname{Max}\{A^{-l}\|T^l\| : 0 \leqq l < n\}$. Hence

$$\|e^{tT}\| = \left\| \sum_0^\infty \frac{t^m T^m}{m!} \right\| \leqq \sum_0^\infty \frac{|t|^m \|T^m\|}{m!} \leqq Me^{A|t|},$$

so $\|e^{isT}\| = \|e^{i\sigma T}e^{-tT}\| \leqq \|e^{-tT}\| \leqq Me^{A|t|}$, for every $s = \sigma + it$, provided $A > 0$. Let $B(X)$ denote the space of all bounded operators on X. By the Hahn-Banach theorem, there exists a continuous linear functional ϕ on $B(X)$ with $\|\phi\| = 1$ and $\phi(T) = \|T\|$. Put

$$f(s) = \phi(e^{isT}) = \sum_0^\infty \frac{\phi(T^m)}{m!} (is)^m.$$

Then f is an entire function, and $|f(s)| \leqq \|e^{isT}\|$. Thus $|f(\sigma)| \leqq 1$ for all real σ, and $|f(s)| \leqq Me^{A|t|}$ if $A > 0$. By Proposition B′, we find $|f'(0)| \leqq A$, if $A > 0$. If $A = 0$, then f is a polynomial, bounded on the real axis, and hence constant, so $f'(0) = 0$. But $f'(0) = i\phi(T) = i\|T\|$. Thus $\|T\| \leqq A$, i.e., $\|T\|^n \leqq \|T^n\|$. Since the reverse inequality comes free, the proposition is proved.

Consider the following example. Let X be the space of all entire functions f such that $\|f\| = \sup\{|f(\sigma+it)|e^{-|t|} : \sigma, t \text{ real}\} < \infty$. Define the operator D on X by $Df = -if'$. For $f \in X$, we have $\|f\| = \sup\{|f(\sigma)| : \sigma \text{ real}\}$ (to see this, apply the maximum principle for bounded analytic functions in half-planes to $f(s)e^{\pm is}$). Thus Bernstein's inequality B′ is equivalent to the assertion that $\|D\| \leqq 1$. For any entire function f, we have the Cauchy estimates $|f^{(n)}(s)| \leqq n!\, r^{-n}M(s; r)$, where $M(s; r) = \operatorname{Max}\{|f(\zeta)| : |\zeta - s| = r\}$; in particular, for $f \in X$ we have $|f^{(n)}(s)| \leqq n!\, r^{-n}e^{|t|+r}\|f\|$ $(s = \sigma + it)$, whence $\|f^{(n)}\| \leqq n!(e/n)^n\|f\|$, i.e., $\|D^n\| \leqq n!(e/n)^n$. Hence $\lim \sup \|D^n\|^{1/n} \leqq 1$. (To see this, put $a_n = n!(e/n)^n$; then $\lim(a_{n+1}/a_n) = e \lim(n/n+1)^n = 1$, so $\lim a_n^{1/n} = 1$.) If $f(s) = \sin s$, then $\|f\| = 1$ and

$\|D^n f\| = 1$ for every n, so $\|D^n\| \geqq 1$ for every n. Thus $\lim \|D^n\|^{1/n} = \inf \|D^n\|^{1/n} = 1$. Hence $\|D\| \leqq 1$ is equivalent to the assertion that $\|D^n\| = \|D\|^n$ for every positive integer n. Now for any real u, any $f \in X$, $\cdots$

$$e^{iuD}f(s) = \sum_0^\infty \frac{u^n}{n!}(iD)^n f(s) = \sum_0^\infty \frac{f^{(n)}(s)}{n!}u^n = f(s+u),$$

whence $\|e^{iuD}f\| = \|f\|$. Thus $\|e^{iuD}\| = 1$ for all real u, i.e., D is Hermitian. Proposition A$'$ then tells us that $\|D^n\| = \|D\|^n$, and thus implies Proposition B$'$.

If T is a Hermitian operator on an inner product space, it is trivial that T^2 is also Hermitian. However, this need not be true for Hermitian operators on normed linear spaces. In fact, consider the operator D just discussed. Let $f(s) = 1 + i \cos s$. Then

$$e^{iuD^2}f(s) = \sum_0^\infty \frac{(-iu)^n}{n!} f^{(2n)}(s) = 1 + ie^{iu}\cos s.$$

Taking $u = \pi/2$, we have $\|e^{iuD^2}f\| = \|1 - \cos s\| = 2$, while $\|f\| = \sqrt{2}$. Thus $\|e^{iuD^2}\| \geqq \sqrt{2}$ for $u = \pi/2$, so D^2 is not Hermitian. Restricting D to the subspace of X spanned by 1, $\sin s$, and $\cos s$, we obtain even an example of a Hermitian operator on a 3-dimensional space whose square is not Hermitian.

Historical Note: Proposition A$'$ was found independently by A. M. Sinclair [3] and the author. A short proof not making use of Bernstein's inequality was given by Bonsall and Crabb [4]. An example of a Hermitian operator T such that T^2 is not Hermitian, quite similar to the one given above, was found independently by Crabb [5]. Previously, Lumer [6] had given an example, a good deal more complicated, of a Hermitian T such that not every power of T was Hermitian. The study of Hermitian operators is a part of the more general theory of the numerical range of operators. This note is an extract from [7], which gives an introduction to that theory, and some further references; a much more comprehensive survey is to be found in the lecture notes of Bonsall and Duncan [8], which have recently appeared.

References

1. I. Vidav, Eine metrische Kennzeichnung der selbstadjungierten Operatoren, Math. Z., 66 (1956) 121–128.

2. G. Lumer, Semi-inner product spaces, Trans. Amer. Math. Soc., 100 (1961) 29–43.

3. A. M. Sinclair, The norm of a Hermitian element in a Banach algebra, to appear.

4. F. F. Bonsall and M. J. Crabb, The spectral radius of a Hermitian element of a Banach algebra, Bull. London Math. Soc., 2 (1970) 178–180.

5. M. J. Crabb, Some results on the numerical range of an operator, J. London Math. Soc., 2 (1970) 741–745.

6. G. Lumer, Spectral operators, Hermitian operators, and bounded groups, Acta Sci. Math. (Szeged)., 25 (1964) 75–85.

7. A. Browder, States, numerical range, etc. Proc. of the Brown Informal Analysis Seminar, summer 1969 (multilith notes).

8. F. F. Bonsall and J. Duncan, Numerical ranges of operators on normed spaces and of elements of normal algebras, London Math. Soc. Lecture Note Series (2), (1971).

CORRECTION TO "FUNDAMENTA MATHEMATICAE: AN EXAMINATION OF ITS FOUNDING AND SIGNIFICANCE"

SISTER MARY GRACE KUZAWA, Holy Family College

It is stated in this article (vol. 77, May 1970, 485–492) that Julius P. Schauder was among the thirty Professors of the Lwow area who were shot by the Gestapo in July, 1941. This is incorrect. Julius P. Schauder was murdered by the Nazis in September of 1943.

I wish to thank Professors H. M. Schaerf and Henryk Fast for detecting the error.

MATHEMATICAL NOTES

EDITED BY DAVID DRASIN

Manuscripts for this Department should be sent to Robert Gilmer, Department of Mathematics, Florida State University, Tallahassee, FL 32306.

ON FUNCTIONS WITH SUMMABLE DERIVATIVE

CASPER GOFFMAN, Purdue University

The following theorem has been known for a long time, [1]:

A function whose derivative exists everywhere and is summable is absolutely continuous. Our purpose is to give a short transparent proof. For proofs in the literature and discussion the reader may consult [1]–[5].

LEMMA 1. *If f is everywhere differentiable on an interval I and $S \subset I$ has measure 0, then $f(S)$ has measure 0.*

Proof. Let $S \subset I$ be of measure 0. For each n and k, let S_{nk} be the set of points $x \in S$ such that for each interval J of length less than $1/k$, with $x \in J$, we have $\mu[f(J)] < n \cdot \mu(J)$. Then $S = \bigcup_{n,k=1}^{\infty} S_{nk}$. Let $\epsilon > 0$. Cover S_{nk} by nonoverlapping closed intervals, each of length less than $1/k$, the sum of whose lengths is less than ϵ/n. Then $\mu[f(S_{nk})] < \epsilon$. So $\mu[f(S_{nk})] = 0$ and $\mu[f(S)] = 0$.

COROLLARY. *If f is everywhere differentiable on $I = [a, b]$ and $\{I_n\}$ is a sequence of nonoverlapping intervals in I with $\sum_{n=1}^{\infty} \mu(I_n) = \mu(I)$, then $\sum_{n=1}^{\infty} \mu[f(I_n)] \geqq |f(b) - f(a)|$.*

Proof. Since $\bigcup_{n=1}^{\infty} f(I_n) \supset f(I)$, except possibly for a set of measure 0, the result follows from Lemma 1.

LEMMA 2. *If f is everywhere differentiable and f' is summable on I, then for every $[c, d] \subset I$,*

$$\int_c^d |f'(t)|\,dt \geqq |f(d) - f(c)|.$$

Proof. For almost all $x \in [c, d]$,

$$\lim_{h\to 0} \left| \frac{1}{h} \int_x^{x+h} f'(t)dt \right| = \lim_{h\to 0} \frac{1}{h} \mu[f([x, x+h])] = |f'(x)|.$$

Let $\epsilon > 0$. Then for almost every x there is an $h(x) > 0$ such that $0 < h < h(x)$ implies

$$\left| \left| \int_x^{x+h} f'(t)dt \right| - \mu[f([x, x+h])] \right| < \epsilon h,$$

so that

$$\int_x^{x+h} |f'(t)|\,dt > \mu[f([x, x+h])] - \epsilon h.$$

By the Vitali Covering Theorem, there are pairwise disjoint intervals $[c_n, d_n]$, for $n = 1, 2, \cdots$, with $\sum_{n=1}^{\infty} (d_n - c_n) = d - c$ and

$$\int_c^d |f'(t)|\,dt = \sum_{n=1}^{\infty} \int_{c_n}^{d_n} |f'(t)|\,dt$$

$$\geqq \sum_{n=1}^{\infty} \mu(f[c_n, d_n]) - \epsilon(d - c) \geqq |f(d) - f(c)| - \epsilon(d - c)$$

by the Corollary. Since this holds for each $\epsilon > 0$, it follows that $\int_c^d |f'(t)|\,dt \geqq |f(d) - f(c)|$.

THEOREM. *If f' exists everywhere and is summable on an interval I, then f is absolutely continuous on I.*

Proof. For each $\epsilon > 0$ there is a $\delta > 0$ such that if the intervals $[a_i, b_i]$, for $i = 1, \cdots, n$, are pairwise disjoint and if $\sum_{i=1}^{n} (b_i - a_i) < \delta$, then

$$\sum_{i=1}^{n} |f(b_i) - f(a_i)| \leqq \sum_{i=1}^{n} \int_{a_i}^{b_i} |f'(t)|\,dt < \epsilon.$$

References

1. H. Lebesgue, Leçons sur l'Intégration, Paris, 1928, p. 183.

2. I. P. Natanson, Theory of Functions of a Real Variable, translated by L. F. Boron, New York, 1955, p. 266.

3. J. Korevaar, Mathematical Methods, Vol. 1, Academic Press, New York, 1968, p. 418.

4. H. Kestelman, Modern Theories of Integration, New York, 1960, p. 183.

5. W. Rudin, Real and Complex Analysis, McGraw-Hill, New York, 1966, p. 168.

A DUAL OF DILWORTH'S DECOMPOSITION THEOREM

L. MIRSKY, University of Sheffield, England

Let P be a partially ordered set. A subset S of P will be called a **chain** if any two elements in S are comparable; it will be called an **antichain** if no two (distinct) elements in S are comparable. In particular, the empty set is both a chain and an antichain. A chain is said to be **maximal** if it is not a proper subset of any chain. An element x in S is said to be **maximal** if $y \leqq x$ for every element y in S which is comparable with x.

We owe to Dilworth [1] the following well-known and important decomposition theorem:

THEOREM 1. *Let* P *be a partially ordered set and* m *a natural number. If* P *possesses no antichain of cardinal* $m+1$, *then it can be expressed as the union of* m *chains.*

It may be of some interest to note that this statement remains valid if the roles of chains and antichains are interchanged. Thus we have the following result:

THEOREM 2. *Let* P *be a partially ordered set, and* m *a natural number. If* P *possesses no chain of cardinal* $m+1$, *then it can be expressed as the union of* m *antichains.*

Thus, in a formal sense, Theorem 2 may be regarded as a 'dual' of Theorem 1. However, as we shall see, the proof of the dual result is considerably easier than that of Dilworth's original theorem. In particular, to establish Theorem 1 we need first to deal with the case where P is finite (see Tverberg's elegant treatment in [5]) and then extend the conclusion to the general case, say by invoking Rado's selection principle (the details can be found, e.g., in [3]). By contrast, a single induction argument suffices to prove Theorem 2.

When $m=1$, the assertion holds trivially. Let $m \geqq 2$; assume that the assertion holds for $m-1$, and let P be a partially ordered set which has no chain of cardinal $m+1$. The antichain M consisting of all maximal elements in P is clearly non-empty since the maximal element of every maximal chain belongs to M. Further, no chain in P\M has cardinal m. For assume, on the contrary, that

$$x_1 < x_2 < \cdots < x_m, \quad x_k \in \mathrm{P}\backslash\mathrm{M} \qquad (1 \leqq k \leqq m).$$

Then, since this chain has cardinal m, it is maximal and so $x_m \in \mathrm{M}$, which contradicts the relation $x_m \in \mathrm{P}\backslash\mathrm{M}$. Since, then, no chain in P\M has cardinal m, it follows by the induction hypothesis that P\M can be expressed as the union of $m-1$ antichains. Hence P can be expressed as the union of m antichains.

We note an easy consequence of Theorem 2.

COROLLARY. *Let* r, s *be positive integers. Then a partially ordered set of* $rs+1$ *elements possesses a chain of cardinal* $r+1$ *or an antichain of cardinal* $s+1$ *or both.*

If there is no chain of cardinal $r+1$, then the given set P can be expressed as the union of r antichains, which may be assumed to be pairwise disjoint, say $P = A_1 \cup \cdots \cup A_r$. Hence, denoting by $|A|$ the cardinal of A, we have

$$rs + 1 = |A_1| + \cdots + |A_r|.$$

Therefore

$$rs + 1 \leqq r \max |A_i|$$

and so $s+1 \leqq \max|A_i|$, as required. It should be noted that the corollary follows in just the same way from Theorem 1, and also that it is best possible in the sense that $rs+1$ cannot be replaced by rs.

In conclusion, we recall a result of Erdös and Szekeres [2] (see Seidenberg [4] for a very short proof) which is an easy consequence of the corollary: *Each sequence of* $rs+1$ *real terms possesses an increasing subsequence of* $r+1$ *terms or a decreasing subsequence of* $s+1$ *terms or both.* The deduction of this result from the corollary appears to be quite well known (or may be left as an exercise for the reader), and we omit the details.

References

1. R. P. Dilworth, A decomposition theorem for partially ordered sets, Ann. of Math., (2), 51 (1950) 161–166.

2. P. Erdös and G. Szekeres, A combinatorial problem in geometry, Compositio Math., 2 (1935) 463–470.

3. L. Mirsky and Hazel Perfect, Systems of representatives, J. Math. Analysis Appl., 15 (1966) 520–568.

4. A. Seidenberg, A simple proof of a theorem of Erdös and Szekeres, J. London Math. Soc., 34 (1959) 352.

5. H. Tverberg, On Dilworth's decomposition theorem for partially ordered sets, J. Combinatorial Theory, 3 (1967) 305–306.

TOPOLOGIES ON ORDERED SETS

F. W. LOZIER, The Cleveland State University

A recent problem in this MONTHLY [1] asks whether it is possible to topologize the integers in such a way that the connected sets are precisely the sets of consecutive integers. The object of this note is to point out that, for a suitable generalization of "sets of consecutive integers," there is a simple necessary and sufficient condition for any partially ordered set to have such a topology.

Let $\langle P; \leqq \rangle$ be a partially ordered set. For $a, b \in P$ we write aRb if and only if $a<b$ and $\{x \in P \mid a<x<b\} = \varnothing$, or $b<a$ and $\{x \in P \mid b<x<a\} = \varnothing$. We say that $\langle a_1, \cdots, a_n \rangle$ is an R-**chain of length** n connecting a_1 and a_n (n may be 1) if and only if $a_i R a_{i+1}$ for $1 \leqq i < n$; if $a_i \in A \subseteq P$ for each i, we say $\langle a_1, \cdots, a_n \rangle$ is an R-**chain** in A. Finally, we say that $A \subseteq P$ is a **set of consecutive elements** of P if and only if for all $a, b \in A$ there is an R-chain in A connecting a and b.

not open in C, or $N_b = St\, b$, in which case B is not open in C. Thus C is connected. Conversely, suppose C is a nonempty connected subset of P. Choose $a \in C$, let A be the set of all elements of C which can be connected to a by an R-chain in C, and let $B = C - A$. Then $x \in A$ implies $C \cap N_x \subseteq C \cap St\, x \subseteq A$ so that A is open in C, and similarly for B. Therefore, since C is connected, it follows that $A = C$. This completes the proof.

Note that for a given $\langle P; \leqq \rangle$, not every topology satisfying (1) need be of the type constructed in the proof; consider the rationals with the usual topology and order, or the indiscrete space $\{a, b\}$ with $a < b$.

Reference

1. Dan Marcus, Problem 5712, this MONTHLY, 77 (1970) 85.

CUBIC TRIANGLE INEQUALITIES

K. B. STOLARSKY, University of Illinois, Urbana

In a recently published book, an attempt was made to list most of the inequalities concerning a, b, c, the lengths of the sides of a triangle (O. Bottema, R. Z. Djordjević, R. R. Janić, D. S. Mitrinović, and P. M. Vasić, *Geometric Inequalities*, Wolters-Noordhoff Publishing, Groningen 1969, pp. 11–17). Of the 24 listed, 13 (1.1–1.8, 1.15–1.17, 1.19, 1.23) involved symmetric forms of degree $n \leqq 3$. Each of these is a special case of the result of this paper. However, inequalities of degree 4, such as

$$a^4 + b^4 + c^4 \leqq 2(a^2b^2 + a^2c^2 + b^2c^2) \leqq 2(a^4 + b^4 + c^4), \tag{1}$$

seem harder to classify. It is curious that (1) is not listed in the book, although it is a simple consequence of the Cauchy-Schwarz inequality, and the identity

$$c^4 + (a^2 - b^2)^2 + (a^2 - (b - c)^2)((b + c)^2 - a^2) = 2c^2(a^2 + b^2). \tag{2}$$

THEOREM. *Let $P(x_1, x_2, x_3)$ be a real symmetric form of degree $n \leqq 3$. If $P(1, 1, 1)$, $P(1, 1, 0)$, and $P(2, 1, 1)$ are all nonnegative, then $P(a, b, c) \geqq 0$. If $P(1, 1, 1) = 0$ and $P(1, 1, 0) > 0$, equality holds if and only if $a = b = c$. However, a real symmetric form of degree 4 can be positive at $(1, 1, 1)$, $(1, 1, 0)$, $(2, 1, 1)$, and $(4, 3, 2)$ while negative at $(1, 1, \frac{1}{2})$.*

Proof. It suffices to prove the first two statements for $n = 3$, since they are much easier when $n = 1, 2$. Since P is symmetric,

$$P(a, b, c) = A(a^3 + b^3 + c^3) + B(a^2b + b^2c + c^2a + ab^2 + bc^2 + ca^2) + Cabc. \tag{3}$$

Define $c_1 = 2A + 2B = P(1, 1, 0)$, $c_2 = 3A + 5B + C$, $c_3 = \frac{3}{2}c_1 + c_2$, $c_6 = 3A + 6B + C = P(1, 1, 1)$, $c_4 = -\frac{3}{2}c_1 - c_2 + 4c_6$, and $c_5 = 2c_6$. By hypothesis, $c_1 \geqq 0$, $c_1 + c_2 = 2P(2, 1, 1) \geqq 0$, and $c_6 \geqq 0$. Hence $c_3 \geqq 0$ and $2c_3 + c_4 = c_3 + 4c_6 \geqq 0$. Now we can assume without loss of generality that $a \geqq b$, and $a \geqq c$. Thus a, b, c will vary over the side lengths of all triangles when $a = x + y + z$, $b = x + y$, $c = y + z$, and x, y, z

vary over all nonnegative real numbers; $a=b=c$ if and only if $x=z=0$. The first two statements now follow from the identity

$$P(a, b, c) = [c_1(x+z)(x-z)^2 + (c_1+c_2)(x^2z+xz^2)] + [c_3(x-z)^2 + (2c_3+c_4)xz]y + [c_5(x+z)]y^2 + c_6y^3. \tag{4}$$

Next, define a symmetric form of degree 4 by

$$P(a, b, c) = A(a^4+b^4+c^4) + B(a^3b+b^3c+c^3a+ab^3+bc^3+ca^3) + C(a^2b^2+b^2c^2+c^2a^2) + D(a^2bc+ab^2c+abc^2), \tag{5}$$

where $A=103/34$, $B=-4$, $C=2$, $D=3$. Then $P(1, 1, 1)=3/34$, $P(1, 1, 0)=1/17$, $P(2, 1, 1)=275/34$, and $P(4, 3, 2)=1815/34$, but $P(1, 1, \frac{1}{2})=-1/(16)(34)$.

Roughly, the theorem asserts that the behavior of P is determined by its behavior at the "most extreme" isosceles triangles.

Problems involving cubic triangle inequalities arise when one tries to improve the case $n=4$ of the following theorem, which for $n=3$ is almost the triangle inequality. Let $|p_ip_j|$ denote the distance between the points p_i and p_j of a metric space S. Let $\prod=\prod(p_1, \cdots, p_n)$ denote the product of all $\binom{n}{2}$ distances determined by $p_1, \cdots, p_n$. Let $\prod_j=\prod_j(p_1, \cdots, p_n)$ denote the product of the $\binom{n-1}{2}$ distances determined by all p_i except p_j.

THEOREM. *If* $n \geqq 3$ *and* $c(n)=2^{-n+1}$, *then* $\sum_{j\neq i}\prod_j \geqq c(n)\prod_i$.

Proof. Without loss of generality, assume all p_i are distinct, $\prod_i$ is maximal among the $\prod_j$, and $\prod_i=\prod_n=\prod(p_1, \cdots, p_{n-1})$. Choose $n_0 \leqq n-1$ so that $|p_np_{n_0}|=\min_{j=1}^{n-1}|p_np_j|$. Then

$$\prod_{n_0} = \prod_i \prod_{j=1}^{n-1} |p_np_j| / |p_{n_0}p_j|,$$

and

$$|p_{n_0}p_j| \leqq |p_{n_0}p_n| + |p_np_j| \leqq 2|p_np_j|.$$

I conjecture that the above is true with $c(n)=(n-1)2^{-n+2}$. For $n=3$ this is simply the triangle inequality. To prove it for $n=4$, set

$$a = |p_2p_3|, \quad b = |p_1p_3|, \quad c = |p_1p_2|,$$
$$a' = |p_1p_4|, \quad b' = |p_2p_4|, \quad \text{and} \quad c' = |p_3p_4|;$$

the assertion becomes

$$\frac{b'c'}{bc} + \frac{c'a'}{ca} + \frac{a'b'}{ab} \geqq \frac{3}{4}.$$

Clearly $b'+c' \geqq a$, $c'+a' \geqq b$, and $a'+b' \geqq c$. Without loss of generality, we may

assume $b'+c'=a$ and $c'+a'=b$ (first shrink a', b', c' simultaneously, then shrink only a'). Thus we must show that

$$Q(c') = c'^2(c - a - b) + c'(a^2 + b^2 - (a + b)c) + \tfrac{1}{4}\, abc \geqq 0.$$

Since $a+b\geqq c$, we have $c'\leqq\frac{1}{2}(a+b-c)$, so it suffices to show that $\frac{1}{2}(a+b-c)$ is at most the largest root of $Q(c')$. A simple calculation reduces this to showing that

$$abc \geqq (c - a + b)(c + a - b)(a + b - c),$$

which follows from our first theorem.

In the case that S is a Euclidean space, the author has some evidence that the above theorem is true with

$$c(n) = (n - 1)2^{-(n-2)/2}(1 + 1/(n - 2))^{-(n-2)/2}.$$

THE SMALLEST SPHERE CONTAINING A RECTIFIABLE CURVE

J. C. C. NITSCHE, University of Minnesota

THEOREM. *Each continuous closed curve of length L in Euclidean 3-space is contained in a (closed) ball of radius $R\leqq L/4$. Equality holds only for a "needle", i.e., a segment of length $L/2$ gone through twice, in opposite directions.*

The estimate $R\leqq L/2$ is obvious. While the theorem appears intuitively clear, the author has not seen it stated in the literature. A simple proof follows:

Proof. Consider a closed curve $\mathfrak{C}$ of length L. Let B be a closed ball of smallest radius containing $\mathfrak{C}$. Choose the coordinate system so that B is defined by the inequality $|\mathfrak{x}| =(x^2+y^2+z^2)^{\frac{1}{2}}\leqq R$. The set of points of $\mathfrak{C}$ on the boundary ∂B—call this set $\mathfrak{C}_0$—must "support" B, i.e., each closed half space $\mathfrak{a}\cdot\mathfrak{x}\geqq 0$ ($\mathfrak{a}$ a constant unit vector) must contain at least one such point. Otherwise B could not have been a ball of smallest radius containing $\mathfrak{C}$.

If two points of $\mathfrak{C}_0$ are diametrical on ∂B then their distance d cannot exceed the value $L/2$ so that $R=d/2\leqq L/4$. Of course, if $R=L/4$, i.e., if $d=L/2$, then $\mathfrak{C}$ must be a needle.

Assume now that there is no pair of diametrical points in $\mathfrak{C}_0$. In this case $\mathfrak{C}_0$ contains at least three points. If $\mathfrak{x}_1$ is one of these points, a second point $\mathfrak{x}_2$ must lie in the half space $\mathfrak{x}\cdot\mathfrak{x}_1\leqq 0$. Choose a coordinate system in which these points are $\mathfrak{x}_1=\{R \sin\alpha, 0, R \cos\alpha\}$ and $\mathfrak{x}_2=\{-R \sin\alpha, 0, R \cos\alpha\}$, where $\pi/4\leqq\alpha<\pi/2$. Still another point $\mathfrak{x}_3=\{R \cos\delta \cos\phi, R \cos\delta \sin\phi, -R \sin\delta\}$ of $\mathfrak{C}_0$ must lie in the half space $z\leqq 0$, so that $0\leqq\delta\leqq\pi/2$ and $0\leqq\phi\leqq 2\pi$. Obviously

$$\begin{aligned} L &\geqq |\mathfrak{x}_1 - \mathfrak{x}_2| + |\mathfrak{x}_2 - \mathfrak{x}_3| + |\mathfrak{x}_3 - \mathfrak{x}_1| \\ &\geqq \sqrt{2}\, R\{\sqrt{2} \sin \alpha + (1 + \cos \alpha \sin \delta - \sin \alpha \cos \delta \cos \phi)^{1/2} \\ &\qquad + (1 + \cos \alpha \sin \delta + \sin \alpha \cos \delta \cos \phi)^{1/2}\}. \end{aligned}$$

If α and δ are fixed, the right hand side is smallest for $\phi=0$ or $\phi=\pi$. Hence

$$\sqrt{2}\,R[\sqrt{2}\sin\alpha+\sqrt{1+\sin(\delta-\alpha)}+\sqrt{1+\sin(\delta+\alpha)}]\leqq L.$$

For fixed α, the left hand side is smallest for $\delta=0$, so that

$$\sqrt{2}\,R[\sqrt{2}\sin\alpha+\sqrt{1-\sin\alpha}+\sqrt{1+\sin\alpha}]\leqq L.$$

Since the expression between the brackets is larger than $2\sqrt{2}$ for $\pi/4\leqq\alpha<\pi/2$, we finally obtain $4R<L$, and the theorem is proved.

Acknowledgment. The preceding work was supported by the Air Force Office of Scientific Research under AFOSR Grant No. 883-67.

A NUMBER FIELD WITHOUT A RELATIVE INTEGRAL BASIS

ROBERT MACKENZIE AND JOHN SCHEUNEMAN, Indiana University

In this note we describe an example of a number field F and a quadratic extension K/F that does not have an integral basis (minimal basis) relative to F.

As usual, denote by $\mathbf{Q}$ the field of rational numbers and by $\mathbf{Z}$ the ring of integers. Let $F=\mathbf{Q}(\gamma)$, where $\gamma^2+14=0$. Let $\mathfrak{O}_F$ be the ring of integers of F. Then $\mathfrak{O}_F=\mathbf{Z}+\mathbf{Z}\gamma$, as is well known. Let $\mathfrak{p}=7\mathfrak{O}_F+\gamma\mathfrak{O}_F$. Then $\mathfrak{p}^2=7\mathfrak{O}_F$, so since $[F:\mathbf{Q}]=2$, $\mathfrak{p}$ is a prime ideal and 7 is ramified in F.

LEMMA. *$\mathfrak{p}$ is not a principal ideal.*

Proof. Suppose $\mathfrak{p}=(a+b\gamma)\mathfrak{O}_F$, with $a, b\in\mathbf{Z}$. Then there are $u, v, x, y\in\mathbf{Z}$ such that $7=(a+b\gamma)\cdot(u+v\gamma)$ and $\gamma=(a+b\gamma)(x+y\gamma)$. Hence

$$\begin{aligned} au-14bv&=7 & ax-14by&=0\\ bu+av&=0 & bx+ay&=1. \end{aligned}$$

Eliminating u, we obtain $-(a^2+14b^2)v=7b$. This implies $b=0$, since otherwise $a^2+14b^2>|7b|$, which would contradict the equation above. It follows that $\mathfrak{p}=a\mathfrak{O}_F$ with $ay=1$. This says $\mathfrak{p}=\mathfrak{O}_F$, a contradiction.

Now let $K=F(\Gamma)$, where $\Gamma^2+7=0$. It is easily checked that $\Delta=\frac{1}{2}(1+\Gamma)$ is an integer in K.

PROPOSITION. *K does not have an integral basis over F.*

Proof. Suppose $\{\mathrm{A}, \mathrm{B}\}$ is an integral basis of K over F. Then there are $\alpha_1, \beta_1, \alpha_2, \beta_2\in\mathfrak{O}_F$ such that $1=\alpha_1\mathrm{A}+\beta_1\mathrm{B}$ and $\Delta=\alpha_2\mathrm{A}+\beta_2\mathrm{B}$. Let $\bar{\mathrm{A}}, \bar{\mathrm{B}}, \bar{\Delta}$ be the conjugates of $\mathrm{A}, \mathrm{B}, \Delta$, respectively, over F. Then

$$\begin{pmatrix}\alpha_1 & \beta_1\\ \alpha_2 & \beta_2\end{pmatrix}\begin{pmatrix}\mathrm{A} & \bar{\mathrm{A}}\\ \mathrm{B} & \bar{\mathrm{B}}\end{pmatrix}=\begin{pmatrix}1 & 1\\ \Delta & \bar{\Delta}\end{pmatrix}.$$

Taking determinants of both sides and squaring gives $\theta^2\delta=-7$, where

$$\theta=\det\begin{pmatrix}\alpha_1 & \beta_1\\ \alpha_2 & \beta_2\end{pmatrix}\in\mathfrak{O}_F \quad\text{and}\quad \delta=\det\begin{pmatrix}\mathrm{A} & \bar{\mathrm{A}}\\ \mathrm{B} & \bar{\mathrm{B}}\end{pmatrix}^2\in\mathfrak{O}_F.$$

Then $\mathfrak{p}^2=7\mathfrak{O}_F=(\theta\mathfrak{O}_F)^2(\delta\mathfrak{O}_F)$, which implies $\theta\mathfrak{O}_F=\mathfrak{O}_F$ and $\mathfrak{p}^2=\delta\mathfrak{O}_F$ by the unique factorization of ideals in $\mathfrak{O}_F$ and the fact that $\mathfrak{p}$ is not principal.

Since θ is a unit, $\{1, \Delta\}$ must be an integral basis for K over F. However. γ/Γ is an integer in K, so $\gamma/\Gamma=\alpha+\beta\Delta$ with α, $\beta\in\mathfrak{O}_F$. Taking conjugates over F, $-\gamma/\Gamma=\alpha+\beta\bar{\Delta}$, so $2\gamma/\Gamma=\beta(\Delta-\bar{\Delta})=\beta\Gamma$, or $2\gamma=-7\beta$. Thus $(2\mathfrak{O}_F)(\gamma\mathfrak{O}_F)=\mathfrak{p}^2(\beta\mathfrak{O}_F)$. But the left side has order 1 at $\mathfrak{p}$ while the right side has order $\geqq 2$ at $\mathfrak{p}$, which is impossible.

THE INJECTIVE ENVELOPE OF THE UPPER TRIANGULAR MATRIX RING

E. E. Bray, K. A. Byrd, and R. L. Bernhardt, University of North Carolina at Greensboro

Given a ring R (with identity 1) and a left R-module M, one calls M an **injective** module if any R-homomorphism from a submodule of a left R-module A into M can be extended to an R-homomorphism of A into M. In particular if M is injective and ϕ is an R-homomorphism from a left ideal I of R into M, then there is an R-homomorphism ψ from R (thought of as a left R-module) into M such that $\psi(a)=\phi(a)$ for every a in I. But $\psi(r)=\psi(r\cdot 1)=r\psi(1)$ for every r in R, so that $\phi(a)=ax$ for every a in I, where $x=\psi(1)$. Thus if M is injective, any R-homomorphism from a left ideal of R into M is simply given by right multiplication by an element of M. It is true that this, in fact, characterizes injectivity. We state this below and refer to Lambek [**2**, p. 88] for a proof.

Baer's Criterion for Injectivity. *A left R-module M is injective if and only if for each left ideal I of R and R-homomorphism $\phi: I\to M$ there is an x in M such that $\phi(a)=ax$ for each a in I.*

If A is a submodule of a left R-module B, we call A an **essential submodule** of B if given any nonzero submodule C of B, we have $A\cap C\neq 0$.

Injective modules are ubiquitous in the sense that each R-module is contained in an injective R-module. This is part of the famous paper by Eckmann and Schopf [**1**] (or Lambek [**2**]) which shows that given an R-module M there is an injective R-module $E(M)$ having M as an essential submodule. This module $E(M)$ is called the **injective envelope** (or **hull**) of M, and it is unique in the sense that any two injective envelopes of M are isomorphic via an R-isomorphism which extends the identity map on M.

The proof of the existence of an injective envelope for a given left R-module depends on Zorn's Lemma; hence the realization of injective envelopes in nature is usually difficult. Perhaps the most readily available one is the injective envelope of Z, the ring of integers treated as a Z-module, for which we have $E(Z)=Q$, the rationals as a Z-module. Another instance which appears in the

The author is grateful to the referee for pointing out Thue's work and for references [1]–[3], and to V. L. Klee for drawing attention to the references in a draft by H. T. Croft and R. K. Guy of their forthcoming book *Research Problems in Intuitive Mathematics*.

Partially supported by Canadian NRC grant No. A3983.

References

1. G. A. Hedlund, Remarks on the work of Axel Thue on sequences, Nordisk Mat. Tidskr., 15 (1967) 147–150. MR 37 (1959), #4454.

2. A. Thue, Über unendliche Zeichenreihen, Norske Vid. Selsk. Skr., I Mat.-Nat. Kl., Christiania 1906, No. 7, 1–22.

3. ———, Über die gegenseitige Lage gleicher Teile gewisser Zeichenreihen, *ibid.*, 1912, No. 1, 1–67.

4. S. Arshon, Démonstration de l'éxistence des suites asymétriques infinies, Mat. Sb. (= Recueil Mat.), 2 (44) (1937) 769–779 (Russian, with French summary).

5. Marston Morse, A solution of the problem of infinite play in chess, Abstract 360, Bull. Am. Math. Soc., 44 (1938) 632.

6. Marston Morse and Gustav A. Hedlund, Unending chess, symbolic dynamics and a problem in semigroups, Duke Math. J., 11 (1944) 1–7. MR 5 (1944) 202.

7. D. Hawkins and W. E. Mientka, On sequences which contain no repetitions, Math. Student, 24 (1956) 185–187. MR 19 (1958) 241.

8. John Leech, A problem on strings of beads, Math. Gazette, 41 (1957) 277–278.

9. C. H. Braunholtz, Solution to Problem 5030 [1962, 439], this MONTHLY, 70 (1963) 675–676.

10. Richard A. Dean, A sequence without repeats on x, x^{-1}, y, y^{-1}, this MONTHLY, 72 (1965) 383–385.

11. H. E. Robbins, On a class of recurrent sequences, Bull. Amer. Math. Soc., 43 (1937) 413–417.

12. G. A. Hedlund and W. H. Gottschalk, A characterization of the Morse minimal set, Proc. Amer. Math. Soc., 15 (1964) 70–74.

13. A. A. Evdokimov, Strongly asymmetric sequences generated by a finite number of symbols, Dokl. Akad. Nauk SSSR, Tom 179 (1968) 1268–1271 or Soviet Math. Dokl., 9 (1968) 536–539.

14. J. Justin, Characterization of the repetitive commutative semigroups, to appear.

15. P. A. B. Pleasants, Non-repetitive sequences, Proc. Camb. Philos. Soc., 68 (1970) 267–274.

16. J. Justin, Généralisation du théorème de van der Waerden sur les semi-groupes répétitifs, to appear in J. of Comb. Theory.

A CONJECTURED CHARACTERIZATION OF CIRCLES

HANS HERDA, Boston State College

I. Suppose C is a simple, closed, rectifiable plane curve having positive perimeter p. Let x be any point on C. With x associate the unique point x' on C whose distance (measured on C) from x is $p/2$. Denote the line segment joining x and x' by s_x and call s_x the *pseudo-diameter of C at x*. Also denote the length of the pseudo-diameter at x by s_x.

The function $f: C \to R^+$ defined by setting $f(x) = s_x$ is continuous with respect to distance on C. Because C is compact, this implies that $s = \min_{x \in C} s_x$ exists. Since C is simple, s is positive.

Now consider the ratio p/s. If C is a circle having positive, finite radius, then $p/s = \pi$. The conjecture is that $p/s \geqq \pi$ for any such curve C, and that $p/s = \pi$ implies C is a circle.

II. Suppose C is again a simple, closed, rectifiable plane curve having positive perimeter p. Let s be defined for C as in **I**. Choose an orientation on C. Now consider all segments t_x of fixed length t, where $0<t<s$, x is a point on C, and t_x joins x to the "next" point $\bar{x}$ on C, i.e., the point for which the arc of C oriented from x to $\bar{x}$ has the same orientation as C, and has minimal length. It can be established that $\bar{x}$ is unique. Denote the length of this arc by a_x.

The function $g: C \to R^+$ defined by setting $g(x) = a_x/t$ is continuous with respect to distance on C. Because C is compact, $r_t = \max_{x \in C} (a_x/t)$ exists. Since a_x is greater than or equal to t for all x, r_t is not less than one.

Now let t vary so that $0<t<s$, and consider the various $r_t \geqq 1$. By the greatest lower bound principle, $\inf_t r_t$ exists. Again, because $r_t \geqq 1$ for all t, it follows that $\inf_t r_t \geqq 1$. The conjecture is that $\inf_t r_t = 1$ if and only if C has a unique tangent at each of its points.

It is natural to try Steiner symmetrization for solving **I**. This approach seems to fail. Both conjectures can be generalized in various ways, but these simple formulations appear most attractive.

CLASSROOM NOTES

Edited by David Drasin

Manuscripts for this Department should be sent to Robert Gilmer, Department of Mathematics, Florida State University, Tallahassee, FL 32306.

AN APPLICATION OF DETERMINANTS

Helen Skala, University of Massachusetts, Boston

A simple and elegant application of the theory of determinants for the beginning student is the following criterion of Sylvester, a well-known theorem of algebraic lore: let K be a field and $f(x) = a_m x^m + \cdots + a_1 x + a_0$, $g(x) = b_n x^n + \cdots + b_1 x + b_0$, where $a_m \neq 0 \neq b_n$, be two polynomials in $K[x]$; then $f(x)$ and $g(x)$ have a nonconstant factor in $K[x]$ if and only if the determinant of the following $(m+n)\times(m+n)$ matrix A is zero:

$$A = \begin{bmatrix} a_m & a_{m-1} & \cdots & a_1 & a_0 & 0 & 0 & \cdots & 0 \\ 0 & a_m & \cdots & a_2 & a_1 & a_0 & 0 & \cdots & 0 \\ \cdot & \cdot & \cdots & \cdot & \cdot & \cdot & \cdot & \cdots & \cdot \\ 0 & 0 & \cdots & & & & & \cdots & a_0 \\ b_n & b_{n-1} & \cdots & & & b_0 & 0 & \cdots & 0 \\ 0 & b_n & \cdots & & & b_1 & b_0 & \cdots & 0 \\ \cdot & \cdot & \cdots & \cdot & \cdot & \cdot & \cdot & \cdots & \cdot \\ 0 & 0 & \cdots & & & & & \cdots & b_0 \end{bmatrix}.$$

We present a simple proof of this theorem which requires only knowledge of

the fact that the determinant of the product of two matrices is the product of the determinants; no use of the theory of linear equations is needed. Set

$$B = \begin{bmatrix} x^{n+m-1} & 0 & 0 & \cdot & \cdot & \cdot & 0 \\ x^{n+m-2} & 1 & 0 & \cdot & \cdot & \cdot & 0 \\ x^{n+m-3} & 0 & 1 & \cdot & \cdot & \cdot & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ x & 0 & 0 & \cdot & \cdot & 1 & 0 \\ 1 & 0 & 0 & \cdot & \cdot & \cdot & 1 \end{bmatrix}.$$

Then $|B| = x^{n+m-1}$ and

$$|AB| = |A|\, x^{n+m-1} = \begin{vmatrix} x^{n-1}f(x) & a_{m-1} & a_{m-2} \cdots 0 \\ x^{n-2}f(x) & a_m & a_{m-1} \cdots 0 \\ \cdot & \cdot & \cdot \quad \cdots \cdot \\ f(x) & \cdot & \cdot \quad \cdots a_0 \\ x^{m-1}g(x) & b_{n-1} & b_{n-2} \cdots 0 \\ \cdot & \cdot & \cdot \quad \cdots \cdot \\ g(x) & \cdot & \cdot \quad \cdots b_0 \end{vmatrix} = f(x)h(x) + g(x)k(x),$$

where $h(x)$ and $k(x)$ are polynomials in $K[x]$ of degree at most $n-1$ and $m-1$, respectively, calculated by expanding $|AB|$ by the first column.

If $f(x)$ and $g(x)$ have a nonconstant factor $r(x)$, then

$$|A|\, x^{n+m-1} = f(x)h(x) + g(x)k(x) = r(x)q(x),$$

for some polynomial $q(x)$. If $q(x)=0$, then clearly $|A|=0$. If $q(x)\neq 0$, then $r(x)$ is a multiple of some power of x. But since $r(x)$ is a factor of both $f(x)$ and $g(x)$, both a_0 and b_0 must be zero, whence the last column of A consists of zeros and again $|A|=0$.

Conversely, suppose $|A|=0$. Then $f(x)h(x)=-g(x)k(x)$. Factoring both sides of this equality into irreducible factors over K we must obtain the same factors, and hence all factors of $f(x)$ must divide either $g(x)$ or $k(x)$. But since $k(x)$ is of at most degree $m-1$, not all factors of $f(x)$ can divide $k(x)$, hence $f(x)$ and $g(x)$ have a common factor.

THE INTERVALS OF CONVERGENCE OF SOME POWER SERIES

EUGENE SCHENKMAN, Purdue University

In this note we deduce the exact radius of convergence for some Maclaurin expansions by methods available to students in a first-year calculus course. We begin by sketching a proof that if $f(x)=\sum_{i=0}^{\infty} a_i x^i$ with $a_0=1$ and $\sum_{i=1}^{\infty} |a_i\lambda^i| < 1$ for some $\lambda > 0$, then the Maclaurin expansion of $1/f(x)$ has a radius of convergence at least λ. For if $0 \leqq |x| \leqq \lambda$, then $f(x) > 0$ and

$$\begin{aligned} \log[f(x)] &= \log\{1 - [1 - f(x)]\} \\ &= -\sum_{n=1}^{\infty} \frac{1}{n}[1 - f(x)]^n = -\sum_{n=1}^{\infty} \frac{1}{n}\left(-\sum_{i=1}^{\infty} a_i x^i\right)^n. \end{aligned}$$

This double series converges absolutely for $|x| \leqq \lambda$, so its terms can be arranged to form a power series $\sum_{j=1}^{\infty} b_j x^j$ which converges to $\log f(x)$ for $|x| \leqq \lambda$. Consequently

$$\frac{1}{f(x)} = e^{-\log f(x)} = \sum_{n=0}^{\infty} \frac{1}{n!}\left(-\sum_{j=1}^{\infty} b_j x^j\right)^n,$$

which can also be written as a power series in x, convergent for $|x| \leqq \lambda$. This power series is the Maclaurin expansion of $1/f(x)$, and its radius of convergence is thus at least λ.

The interval of convergence of the function under consideration will be henceforth denoted by ρ.

THEOREM 1. *The interval of convergence for* $\sec x$ *is* $\rho = \frac{1}{2}\pi$.

Proof. It is clear that $\rho < \frac{1}{2}\pi$ since $\sec \frac{1}{2}\pi = \infty$.

In the following we shall need to know that all the coefficients of the Maclaurin expansion of $\sec x$ are nonnegative. To see this we show inductively that the nth derivative of $\sec x$ is $\sec x$ times a polynomial $P_n(\tan x)$ in $\tan x$ with nonnegative coefficients: Indeed, $d(\sec x)/dx = \sec x \tan x$, and if $d^n(\sec x)/dx^n$ were $\sec x\, P_n(\tan x)$, then

$$\begin{aligned}\frac{d^{n+1}\sec x}{dx^{n+1}} &= \sec x[P_n'(\tan x)(\tan^2 x + 1) + P_n(\tan x)\tan x] \\ &= \sec x P_{n+1}(\tan x),\end{aligned}$$

where P_{n+1} is a polynomial with nonnegative coefficients. Now

$$\begin{aligned}\sec x = \frac{1}{\cos x} = \frac{1}{2\cos^2(\frac{1}{2}x) - 1} &= \tfrac{1}{2}\sec^2(\tfrac{1}{2}x)\frac{1}{1 - \frac{1}{2}\sec^2(\frac{1}{2}x)} \\ &= \tfrac{1}{2}\sec^2(\tfrac{1}{2}x)\sum_{n=0}^{\infty}[\tfrac{1}{2}\sec^2(\tfrac{1}{2}x)]^n,\end{aligned}$$

with the last equality valid if $\frac{1}{2}\sec^2(\frac{1}{2}x) < 1$, that is, $\frac{1}{2}x < \frac{1}{4}\pi$. The last sum can be written as a double sum in terms of the expansion of $\sec(\frac{1}{2}x)$, and since this expansion has non-negative terms, the whole expression on the right is a power series in $\frac{1}{2}x$. Hence from the validity of the expansion of secant for $\frac{1}{2}x < \lambda$, follows its validity for $x < 2\lambda$ as long as $\lambda < \frac{1}{4}\pi$, so $\rho = \frac{1}{2}\pi$.

THEOREM 2. *The interval of convergence for* $\tan x$ *and* $\operatorname{sech} x$ *is* $\rho = \frac{1}{2}\pi$.

This is clear since $\tan x = \sin x \sec x$, and $\operatorname{sech} x = \sec ix$.

THEOREM 3. *The interval of convergence for* $x/\sinh x$ *is* $\rho = \pi$.

Proof. Since $\sinh x = 2\sinh(\frac{1}{2}x)\cosh(\frac{1}{2}x)$, it follows that

$$\frac{x}{\sinh x} = \operatorname{sech}(\tfrac{1}{2}x)\frac{\frac{1}{2}x}{\sinh(\frac{1}{2}x)}\cdot$$

Thus the expansion of $x/\sinh x$ is valid for $\frac{1}{2}x < \lambda$ provided $\lambda < \frac{1}{2}\pi$. Hence $\rho = \pi$.

THEOREM 4. *The interval of convergence for $x/(e^x - 1)$ is $\rho = 2\pi$.*

Proof. We have

$$\frac{e^x - 1}{x} = e^{x/2}\,\frac{e^{x/2} - e^{-x/2}}{x} = e^{x/2}\,\frac{\sinh(\frac{1}{2}x)}{\frac{1}{2}x},$$

hence

$$\frac{x}{e^x - 1} = \frac{\frac{1}{2}x}{\sinh(\frac{1}{2}x)}\,e^{-x/2}.$$

It follows that $\rho = 2\pi$. Cf. [1, p. 139].

The author is grateful to Professor G. R. MacLane for some illuminating discussions during the preparation of this note.

Reference

1. I. I. Hirschman, Jr., Infinite Series, Holt, Rinehart and Winston, New York, 1962.

METRICALLY ISOLATED SETS

C. C. ALEXANDER, University of Virginia

A well-known property of compact subsets of a metric space is that of being distant from each disjoint closed set. However, noncompact subsets of metric spaces may also have this property, as may be seen by considering any infinite discrete metric space. However, that there is a relationship between this property and compactness will be illustrated explicitly in this paper. The author wishes to thank the referee for his suggestions and particularly for the simplification obtained by the addition of the corollary to Theorem 2.

Since there is more than one way to view the property mentioned above, we shall be interested in considering sets satisfying the following definitions. Let A be a subset of a metrizable space X. Then A is **metrically isolated** if there is a compatible metric d for X such that $d(A, B) > 0$ for each closed set $B \subset X \backslash A$; it is **absolutely metrically isolated** if for each compatible metric d on X, we have $d(A, B) > 0$ for each closed set $B \subset X \backslash A$. It is clear that a metrically isolated set is closed, and that a compact subset of a metrizable space is absolutely metrically isolated.

For all definitions and notation not specifically given in this paper, we refer the reader to [1]. For a set A, $\mathrm{cl}(A)$ and $\mathrm{Fr}(A)$ will denote the closure and boundary of A, respectively. If d is a metric for X, $x \in X$, and $\epsilon > 0$, then $B(x; \epsilon) = \{y \in X \mid d(x, y) < \epsilon\}$.

At the outset, we have the following pleasant characterization of metrically isolated subsets:

THEOREM 1. *Let A be a subset of a metrizable space X. Then A is a metrically isolated set if and only if A is a closed set with compact boundary.*

Proof. For the necessity, assume A is metrically isolated. It has already been mentioned that A must be a closed set. Let d be a compatible metric for X such that $d(A, B)>0$ for each closed set $B\subset X\backslash A$. Assume $\mathrm{Fr}(A)$ is not compact. Then there is a sequence $\{a_n\}_1^\infty$ in $\mathrm{Fr}(A)$ with no accumulation point. Since $a_n\in\mathrm{Fr}(A)$ for each n, for each n there is $x_n\in X\backslash A$ such that $d(x_n, a_n)<1/n$. Then the set $B=\{x_n \mid n=1, 2, \cdots\}$ has no cluster points, since any cluster point of B is an accumulation point of the sequence $\{a_n \mid n=1, 2, \cdots\}$. Therefore B is a closed subset of X disjoint from A for which $d(A, B)=0$. This is a contradiction. Hence $\mathrm{Fr}(A)$ is compact.

For the sufficiency, assume that A is a closed set with compact boundary. Then the natural quotient map $f:X\to Y=X/A$ is a closed continuous map for which $\mathrm{Fr}(f^{-1}(y))$ is compact for each $y\in Y$. Hence X/A is metrizable, by the Morita-Hanai-Stone Theorem [**1**; p. 254; #15]. Let d and ρ be compatible metrics for X and X/A, respectively. Then it is easily checked that $e(x_1, x_2)=d(x_1, x_2)+\rho(f(x_1), f(x_2))$ defines a compatible metric for X which satisfies the required condition.

COROLLARY. *A necessary and sufficient condition that a metrizable space X be metrically isolated in each metric space in which it is embedded is that X be compact.*

Proof. The sufficiency of the condition is well known. For the necessity, consider $X\simeq X\times\{0\}\subset X\times E^1$ (E^1=real numbers with usual topology). Since $X\times\{0\}$ is metrically isolated in $X\times E^1$ by assumption, $\mathrm{Fr}(X\times\{0\})=X\times\{0\}$ is compact by Theorem **1** and therefore so is X.

We now consider absolutely metrically isolated sets in metrizable spaces, and for this we need the following result; the construction used is essentially similar to that given by Niemytzki and Tychonoff in [**2**].

THEOREM 2. *Let X be a metrizable space and let $\{W_n\}_1^\infty$ be a sequence of open sets in X such that* $\mathrm{cl}(W_{n+1})\subset W_n$ *for each n and $\bigcap\{W_n \mid n=1, 2, \cdots\}=\varnothing$. Then there is a metrizable space X^* such that*

(i) $X^*=X\cup\{\hat{x}\}$, *where* $\hat{x}\notin X$,

(ii) *X is a subspace of X^*,*

(iii) $\{W_n^* \mid n=1, 2, \cdots\}$ *is a local base at $\hat{x}$, where* $W_n^*=W_n\cup\{\hat{x}\}$ *for each n.*

Proof. Choose $\hat{x}\notin X$ and define a topology for $X^*=X\cup\{\hat{x}\}$ having as open sets:

(a) any open set of X, considered as a subset of X^*, and

(b) any set of the form $U\cup\{\hat{x}\}$, where U is an open set in X and $W_n\subset U$ for some n.

It is easily verified that this defines a topology for X^* so that (i), (ii), and (iii) are satisfied. That X^* is regular at $\hat{x}$ follows from the condition on the sequence $\{W_n\}_1^\infty$ that $\mathrm{cl}(W_{n+1})\subset W_n$ for each n. It is then easily seen that X is a

regular Hausdorff space. By the Nagata-Smirnov-Bing Metrization Theorem [1; p. 194], since X is metrizable it has a σ-locally finite basis $\sum = \bigcup\{\sum_n \mid n=1, 2, \cdots\}$. For each pair of positive integers n, m, let $\sum_{m,n} = \{S\backslash \mathrm{cl}(W_n) \mid S\in\sum_m\}$. It follows that $\bigcup\{\sum_{m,n} \mid m, n=1, 2, \cdots\}\cup\{W_n^* \mid n=1, 2, \cdots\}$ is a σ-locally finite basis for X^*. Hence X^* is metrizable.

COROLLARY. *If $\{x_n \mid n=1, 2, \cdots\}$ is a countably infinite closed discrete subset of a metrizable space, then there is a compatible metric d for X such that the sequence $\{x_n\}_1^\infty$ is d-Cauchy.*

Proof. Let ρ be any compatible metric for X. We can assume without loss of generality that $x_n\neq x_m$ for $n\neq m$. For each n, let

$$W_n = \bigcup\{B(x_k; 1/n) \mid k = n, n+1, \cdots\}.$$

The sequence $\{W_n\}_1^\infty$ is easily seen to satisfy the conditions of Theorem 3, since any point in $\bigcap\{W_n \mid n=1, 2, \cdots\}$ must be a cluster point of the set $\{x_n \mid n=1, 2, \cdots\}$, and by assumption, this set has no cluster points. Thus, by Theorem 2, we can add a point $\hat{x}$ to X to obtain a space X^* as described in the theorem. Let d^* be any compatible metric for X^*, and let d be the restriction of d^* to the subspace X of X^*. Since $x_n\in W_n\subset W_n^*$ for each n, it follows that $\{x_n\}_1^\infty$ converges to $\hat{x}$ in X^*. Thus $\{x_n\}_1^\infty$ is d^*-Cauchy and consequently is d-Cauchy, so the theorem is proved.

REMARK. We note that if $\{x_n\}_1^\infty$ is a sequence in X which has no accumulation point, then the *set* $\{x_n \mid n=1, 2, \cdots\}$ is a countably infinite closed discrete subset of X. Since, in metric spaces, compactness is equivalent to the existence of accumulation points for every sequence, it follows from the corollary that each noncompact metric space has a compatible metric on it which is not complete. Thus what we have done in Theorem 2 and its corollary is to break the Niemytzki-Tychonoff proof of the sufficiency in the following theorem into two parts. The theorem is due to Niemytzki and Tychonoff [2]. It is included here, but is not needed for any proofs in this paper.

THEOREM 3 (Niemytzki-Tychonoff). *A metrizable space is compact if and only if it is complete with respect to every compatible metric on it.*

We are now able to obtain the following simple characterization of absolutely metrically isolated sets in a metrizable space.

THEOREM 4. *Let A be a subset of the metrizable space X. Then A is absolutely metrically isolated if and only if A is closed and either A or* $\mathrm{cl}(X\backslash A)$ *is compact.*

Proof. That the condition is sufficient is easily seen; for let A be a closed set such that either A or $\mathrm{cl}(X\backslash A)$ is compact, let B be a closed set disjoint from A, and let d be any compatible metric for X. Then, since A and B are disjoint closed sets, one of which is compact, $d(A, B)>0$. Hence A is absolutely metrically isolated.

Now assume that A is absolutely metrically isolated. Then A is a closed set with compact boundary, by Theorem 1. Suppose that neither A nor $\mathrm{cl}(X\backslash A)$ is compact. Then, as noted above, there is a countably infinite closed (in A) discrete subset $B=\{b_n \mid n=1, 2, \cdots\}$ of A. Since A is closed in X, it follows that B is also closed in X. Similarly there is a closed (in X) discrete subset $C=\{c_n \mid n=1, 2, \cdots\}$ of $\mathrm{cl}(X\backslash A)$. Since $\mathrm{Fr}(A)$ is compact, $C\cap \mathrm{Fr}(A)$ is finite; therefore we may assume, without loss of generality, that $C\cap \mathrm{Fr}(A)=\varnothing$ and consequently that $C\cap A=\varnothing$. Since $B\cup C$ is a closed discrete subset of X, it follows from the corollary to Theorem 2 that there is a compatible metric d for X such that the sequence $\{z_n\}_1^\infty$ is d-Cauchy, where

$$z_n = \begin{cases} b_k & \text{if } n = 2k-1 \\ c_k & \text{if } n = 2k. \end{cases}$$

Then $d(A, C)\leqq d(B, C)=0$, and C is a closed set disjoint from A. Hence A is not absolutely metrically isolated.

References

1. J. Dugundji, Topology, Allyn and Bacon, Boston, Mass. 1965.
2. V. Niemytzki and A. Tychonoff, Beweis des Satzes, dass ein metrisierbarer Raum dann und nur dann kompakt ist wenn er in jeder Metrik vollständig ist, Fund. Math., 12 (1928) 118–120.

THE NON-EXISTENCE OF A BANACH SPACE OF COUNTABLY INFINITE HAMEL DIMENSION

W. R. BAUER, Department of Defense, and R. H. BENNER, IBM Corporation, Washington

The cardinality of the algebraic (Hamel) basis of an infinite dimensional Banach space is greater than $\aleph_0$ (the first countable ordinal). This paper presents an elementary proof which does not use the category theory. Our proof can be presented near the beginning of a course on infinite dimensional spaces at a time when the question arises naturally. Textbooks either omit the result or defer it until after the category theorem is proved.

We recall the following lemma, due to F. Riesz. The lemma is proved in Taylor [1] using only elementary arguments.

LEMMA. *Suppose X is a normed linear space and that X_0 is a closed proper subspace of X. Then for each θ such that $0<\theta<1$ there exists a vector $x_\theta\in X$ such that $\|x_\theta\|=1$ and $\|x-x_\theta\|>\theta$ for all $x\in X_0$.*

THEOREM. *If X is an infinite dimensional Banach space, then X has dimension strictly greater than $\aleph_0$.*

Proof. Suppose not, and let $B=\{z_1, z_2, \cdots\}$ be a countable algebraic basis for X. A new algebraic basis is constructed as follows. Let $x_1=\|z_1\|^{-1}z_1$, and assuming $x_1, \cdots, x_n$ are defined, let $x_{n+1}\in \mathrm{sp}\{z_1, \cdots, z_{n+1}\}$ be of norm 1, and

be such that $\|x_{n+1}-x\|>1/2$ for all $x\in\text{sp}\{x_1, \cdots, x_n\}$. Riesz's lemma insures the existence of x_{n+1}, and since $\text{sp}\{x_1, \cdots, x_n\}=\text{sp}\{z_1, \cdots, z_n\}$ for each n, we do obtain a basis.

Let $y=\sum_{i=1}^{\infty}4^{-i}x_i$. The partial sums form a Cauchy sequence, hence $y\in X$. Since $\{x_1, x_2, \cdots\}$ is an algebraic basis, we may represent y as a finite linear combination, $y=\sum_{i=1}^{n}\alpha_i x_i$. Consequently

$$y - y = \sum_{i=1}^{n}(\alpha_i - 4^{-i})x_i - 4^{-(n+1)}x_{n+1} - \sum_{i=n+2}^{\infty}4^{-i}x_i = 0.$$

Hence

$$\left\|\sum_{i=1}^{n}(\alpha_i - 4^{-i})x_i - 4^{-(n+1)}x_{n+1}\right\| = \left\|\sum_{i=n+2}^{\infty}4^{-i}x_i\right\|,$$

or

$$4^{-(n+1)}\left\|\sum_{i=1}^{n}4^{n+1}(\alpha_i - 4^{-i})x_i - x_{n+1}\right\| = \left\|\sum_{i=n+2}^{\infty}4^{-i}x_i\right\|.$$

Since $\sum_{i=1}^{n}4^{n+1}(\alpha_i-4^{-i})x_i\in\text{sp}\{x_1, \cdots, x_n\}$, the left hand side of the equation must be at least $4^{-(n+1)}\times 1/2$. We then obtain the contradiction that

$$\frac{1}{2}\times 4^{-(n+1)} \leqq \left\|\sum_{i=n+2}^{\infty}4^{-i}x_i\right\| \leqq \sum_{i=n+2}^{\infty}4^{-i}\|x_i\| = \sum_{i=n+2}^{\infty}4^{-i} = \frac{1}{3}4^{-(n+1)}.$$

The contradiction leads us to conclude that an infinite dimensional Banach space cannot have a countable algebraic basis.

References

1. A. E. Taylor, Introduction to Functional Analysis, Wiley, New York, 1958.
2. A. Wilansky, Functional Analysis, Blaisdell, New York, 1964.

ON LINEAR DIFFERENTIAL EQUATIONS WITH CONSTANT COEFFICIENTS

A. K. Bose, Clemson University

1. Introduction. Consider an nth order normalized linear homogeneous differential equation with constant coefficients

$$y^{(n)} + a_{n-1}y^{(n-1)} + \cdots + a_0 y = 0. \tag{1}$$

From the definition of a solution of (1) on the complex plane C, it follows that every solution of (1) on C is analytic on C. Also, it is well known from established results that the solution space S of (1) on C has the following properties:

(a) S is a finite dimensional subspace ($\dim S=n$) of the linear space of all functions analytic on C.

(b) The initial value problem:

$$y^{(n)} + a_{n-1}y^{(n-1)} + \cdots + a_0 y = 0,$$
$$y(z_0) = b_1, \quad y'(z_0) = b_2, \cdots, y^{(n-1)}(z_0) = b_n, \quad z_0 \in C,$$

for any arbitrary vector $(b_1, b_2, \cdots, b_n)$, has a unique solution in S.

(c) $f \in S$ implies that $f' \in S$.

(d) There is a nonnegative integer p such that no member f of S can have a zero of order $p+1$ unless $f\equiv 0$. (That is, $f(z_0) = f'(z_0) = \cdots = f^{(p)}(z_0) = 0$, for some f in S and some $z_0 \in C$ implies $f \equiv 0$.)

(e) The smallest nonnegative integer p for which the property (d) is satisfied is $n-1$.

(f) S is not the solution space of any normalized homogeneous differential equation with constant coefficients other than (1).

The properties (d) and (e) are consequences of property (b) and follow from the following facts:

(i) The unique solution of (1) satisfying the initial condition

$$y(z_0) = y'(z_0) = \cdots = y^{(n-1)}(z_0) = 0,$$

$z_0 \in C$, is $f \equiv 0$. Hence, the property (d) is satisfied for all integers $p \geqq n-1$.

(ii) The unique solution g of (1) satisfying the initial conditions

$$y(0) = y'(0) = \cdots = y^{(n-2)}(0) = 0, \qquad y^{(n-1)}(0) = 1$$

is nontrivial. That is, S contains a nontrivial member which has a zero of order $n-1$.

The purpose of this note is to show that properties (c) and (d) together imply properties (a), (b), (e), and (f), and characterize solution spaces on C of linear homogeneous differential equations with constant coefficients in the sense of the following theorem:

THEOREM 1. *Let S be a subspace of the linear space of all complex-valued functions analytic on C such that $f \in S$ implies $f' \in S$, and there is a nonnegative integer p such that no member f of S can have a zero of order $p+1$ unless $f \equiv 0$. Then*

(i) *S is finite dimensional and* $\dim S = n$, *where $n-1$ is the smallest integer p satisfying the hypothesis,*

(ii) *S is the solution space, on C, of a unique normalized linear homogeneous differential equation of order n with constant coefficients.*

2. *Proof of Theorem* 1. Let $n-1$ $(n \geqq 1)$ be the smallest nonnegative integer p satisfying the hypotheses and let $\{f_j\}_{j=1}^{n+1}$ be any $n+1$ elements of S. Since the system of equations

$$\sum_{j=1}^{n+1} c_j f_j^{(k)}(0) = 0, \quad k = 0, 1, 2, \cdots, n-1,$$

has always a nontrivial solution $(c_1, c_2, \cdots, c_{n+1})$, it follows that the element $g=\sum_{j=1}^{n+1} c_j f_j$ of S has a zero of order n. Hence $g\equiv 0$. This implies that any $n+1$ elements of S are linearly dependent on C. Therefore S is finite dimensional and dim $S\leqq n$. Again there is a nontrivial (distinguished) member h of S which has a zero of order $n-1$. That is, there is a $z_0\in C$ such that $h(z_0)=h'(z_0)=\cdots=h^{(n-2)}(z_0)=0$, $h^{(n-1)}(z_0)\neq 0$. We claim that $\{h, h', \cdots, h^{(n-1)}\}$ is a set of n linearly independent members of S. For suppose that

$$\sum_{j=0}^{n-1} c_j h^{(j)} \equiv 0, \quad \text{on } C. \tag{2}$$

Then,

$$\sum_{j=0}^{n-1} c_j h^{(j+k)}(z_0) = 0, \quad \text{for } k = 0, 1, 2, \cdots. \tag{3}$$

Taking $k=0, 1, 2, \cdots, n-1$ successively in (3), we see that $c_0=c_1=\cdots=c_{n-1}=0$, which proves the linear independence of $\{h, h', \cdots, h^{n-1}\}$. Hence dim $S=n$. This proves conclusion (i).

Since $h^{(n)}\in S$, there exist constants $a_{n-1}, a_{n-2}, \cdots, a_0$ such that

$$h^{(n)} = \sum_{j=1}^{n} a_{n-j} h^{(n-j)}. \tag{4}$$

To prove conclusion (ii), we need to exhibit an nth order normalized linear homogeneous differential equation with constant coefficients whose solution space on C is S. Our candidate for such an equation is

$$y^{(n)} = \sum_{j=1}^{n} a_{n-j} y^{(n-j)}. \tag{5}$$

To support our claim, let $f=\sum_{i=0}^{n-1} c_i h^{(i)}$ be an arbitrary member of S. Then

$$\begin{aligned}\sum_{j=1}^{n} a_{n-j} f^{(n-j)} &= \sum_{j=1}^{n} a_{n-j}\left(\sum_{i=0}^{n-1} c_i h^{(n+i-j)}\right)\\ &= \sum_{i=0}^{n-1} c_i\left(\sum_{j=1}^{n} a_{n-j} h^{(n-j+i)}\right)\\ &= \sum_{i=0}^{n-1} c_i h^{(n+i)} = f^{(n)},\end{aligned}$$

which proves that S is a subspace of the solution space S^* of (5) on C. Since S^* satisfies the hypotheses above with $p=n-1$, it follows, by the above argument, that dim $S^*\leqq n$. Finally, dim $S^*\leqq n$, dim $S=n$, and $S\subset S^*$ imply that $S=S^*$. Hence S is the solution space on C of an nth order normalized linear homogeneous differential equation with constant coefficients. To prove the uniqueness

of (5), suppose that S is the solution space on C of another normalized equation with constant coefficients

$$y^{(m)} = \sum_{j=1}^{m} b_{m-j} y^{(m-j)}. \tag{6}$$

The integer m cannot be less than n, otherwise the distinguished member h of S will not be a solution of (6). So let $m > n$. Consider the characteristic equations of (5) and (6):

$$P_1(x) = x^n - \sum_{j=1}^{n} a_{n-j} x^{n-j} = 0, \tag{7}$$

$$P_2(x) = x^m - \sum_{j=1}^{m} b_{m-j} x^{m-j} = 0. \tag{8}$$

Since the degree of P_2 is greater than that of P_1, either

(a) there is a root α of (8) which is not a root of (7), or

(b) there is a common root of each of (7) and (8) with different multiplicities. In case (a), $e^{\alpha z}$ is a solution of (8) but is not a solution of (7). In case (b), let β be a common root of each of (7) and (8) with multiplicities n_1 and n_2 respectively such that $n_1 \neq n_2$. Without loss of generality we can assume that $n_2 > n_1$. Then $z^{n_2-1}e^{\beta z}$ is a solution of (8) but is not a solution of (7). This contradicts the fact that S is the common solution space of each of (5) and (6). Therefore $m = n$. Again the distinguished member h of S must satisfy each of the equations (5) and (6) (with $m = n$). That is,

$$h^{(n)} = \sum_{j=1}^{n} a_{n-j} h^{(n-j)},$$

$$h^{(n)} = \sum_{j=1}^{n} b_{n-j} h^{(n-j)}.$$

Since $h, h', \cdots, h^{(n-1)}$ are linearly independent on C, it follows that

$$a_{n-j} = b_{n-j}, \qquad j = 1, 2, \cdots, n.$$

This completes the uniqueness proof.

REMARK. It is well known that every solution, on the real line R, of a normalized linear homogeneous differential equation with constant coefficients is analytic on R. Hence the hypotheses (i) and (ii) of Theorem 1 also characterize solution spaces, on the real line R, of linear homogeneous differential equations with constant coefficients.

MATHEMATICAL EDUCATION

EDITED BY J. G. HARVEY AND M. W. POWNALL

Material for this Department should be sent to either of the editors: J. G. Harvey, Department of Mathematics, University of Wisconsin, Madison, WI 53706; M. W. Pownall, Department of Mathematics, Colgate University, Hamilton, NY 13346.

NEW APPROACHES TO GENERAL EDUCATION MATHEMATICS FOR DEVELOPING COLLEGES

BEAUREGARD STUBBLEFIELD, Institute for Services to Education, and
CARL WHITMAN, Florida A. & M. University

1. Introduction. In most mathematics departments, freshman general education mathematics courses affect more college students than all of the other mathematics offerings. Yet, the general education courses are of little interest to these departments, which often relegate them to graduate students. In such courses, many developing colleges are finding "rough diamonds"—able but reluctant or handicapped students—who are potential "late bloomers." They are finding that effective teaching can release mental fetters by breaking away from excessive formalism and stereotyped thinking. Furthermore, they find that as students experience intellectual success they raise their self-expectations and travel new routes to higher achievement. A most effective way of improving student experiences in freshman courses is through programs which involve the instructors in creating fresh approaches to their teaching. They are setting up active learning situations by using intrinsically interesting materials, by providing students with attainable goals, and by allowing flexibility in course content.

The Thirteen College Curriculum Program (T.C.C.P.) is a consortium of developing colleges which aims to improve freshman instruction and curriculum materials. As a large and promising project it is supported by private and public funds. It was conceived with inputs and guidance from such people as Jerold Zacharais and Phillip Morrison of M.I.T., Walter Talbott of Morgan State College, Herman Branson of Howard University, and other distinguished educators. The Institute for Services to Education (I.S.E.) of 2001 S Street, N.W., Washington, D.C. (Dr. Elias Blake, President) provided expert guidance and know-how. The Program was launched in the Summer of 1967 with a writing conference. The conferees devised a new freshman program which attempted to release students from intellectual ruts in formalism and boredom. The course was called "Quantitative and Analytical Thinking," and the materials and techniques were tested on the thirteen campuses the following academic year. (Participants worked in close liaison with curriculum experts of the Curriculum Resources Group of I.S.E. who provided much of the inspiration for the emergent Thirteen College philosophy and techniques.) This pattern was repeated in successive years.

2. Conditions that motivated the development of the T.C.C.P. All colleges and universities have, to some extent, the same educational problems that gave rise to the T.C.C.P. Many of these problems are more acute on black campuses because the colleges have had to "make do" for years with inadequate funding. Moreover, many students have poor academic backgrounds, and teaching modes have been heavily teacher centered. For this reason, perhaps, the attempt to find answers to correct these ills has started here.

The educational problems with which T.C.C.P. has become involved seem to stem from *institutional* or *teacher-centered* conveniences. Many students feel that their college or university fails to meet their needs. They cry out for meaningfulness, accountability, and relevance in their education. The T.C.C.P. has tried to grapple with many facets of these problems and especially with the following deleterious *institutional-teacher-centered* practices:

(1) *The pacing of a group of students using a class-average rate*—this is too fast for some and too slow for others.

(2) *The imposition of a normal distribution on class grading*—this usually means that 10 to 30 percent of the students will fail.

(3) *The forcing of students to be docile*—initiative always belongs to the professor; many gifted and creative students rebel by dropping out of school while conformists remain.

(4) *The mass handling of students in large classes*—this works to the disadvantage of the student; it is done to gain freedom to pursue research and to teach small advanced classes.

(5) *The presentation of mathematics as tightly packaged deductive units*—this approach is convenient for the teacher; it restricts exploratory questions and discourages intuitive jumps; it requires only the time-worn aids of textbook, blackboard, and lecture.

3. The Thirteen College Curriculum Conferences. Summer conferences of eight weeks' duration have provided the primary means for developing motivation and enthusiasm for curriculum innovation among the participating college teachers. They have undertaken to create strategies and materials that would provide effective, meaningful learnings.

The conferences have emphasized several pedagogical principles including:

(1) Recognizing students' needs by involving them in active learning and setting attainable goals.

(2) Using intrinsically interesting puzzles, games, apparatus, and other materials for student investigations that pay off with induced mathematical generalizations, extended intuitions, and new learning strategies.

(3) Recognizing that continuing success motivates students and builds a healthy self-concept. This permits a student to focus on what he can do, without emphasizing the areas in which he may have performed poorly.

(4) Being flexible so that students and teachers can pursue their strengths and interests. Courses on different campuses are not identical. Beyond

minimal requirements, students express their own preferences by pursuing special investigations.

(5) Increasing the instructor's sensitivity to students, his skill in fostering learning, and his expertise in his subject.

To implement these principles, the Thirteen College summer conferences provided on-site experimental classes with college students and seminars to study materials. These helped the participants to design and write descriptions of teaching strategies.

Several types of classroom management are typical of T.C.C.P. approaches. For example, the instructor may pose a problem such as "How many different lines can be drawn between pairs of dots in a circle of 100 dots?" If the instructor poses the problem at the beginning of a class, he circulates among the students speaking only to clarify the problem or to motivate a student. When answers are not forthcoming, the students are encouraged to grapple with the problem working singly or in groups and to share information as they progress toward a solution. Another typical approach is to have students pursue a variety of investigations, such as actually permuting a set of colored cubes. Students then make observations and sense patterns, which intuitively lead to mathematical generalizations and formulations. Another typical technique is for a teacher to lead and guide class discussions but refuse to give any answers. The students understand that only one of the answers is correct. When the teacher refuses to select the correct answer, the class undertakes more thoughtful considerations until some student presents an argument which convinces the others that his answer is the correct one.

4. Thirteen College Experimentation and Observations. Experiments using different formats and procedures on the several campuses have tested many student-centered hypotheses including the following:

(1) Grades for freshman courses should be de-emphasized and, if possible, eliminated.

(2) Class discussions should follow student interests as expressed in class.

(3) Motivation should be intrinsic in the academic work.

(4) Topics should be chosen by student relevance and not by subject structure.

(5) Physical apparatus promotes learning transferable outside of class and provides valuable entree to many mathematical topics.

(6) Students should be encouraged to exercise initiative and should be active in class.

Results from standardized tests have indicated that students in the experimental sections did not suffer academically but even improved slightly when compared to those in traditional sections. An important gain established by special tests was that students improved their self-concept. A further gain from the experimental techniques and materials is that students considered their mathematics classes challenging and enjoyable.

Subjective observations and generalizations by project participants indicate the following:

(1) *When students exercise initiative and are active, excitement and learning increase.* A noisy classroom was not antithetical to learning as long as work and discussion centered on mathematics. Student-student discussions promoted learning in decentralized classrooms.

(2) *Physical apparatus stimulates realistic thinking.* Apparatus used in laboratory experiments generated questions and learnings related to the day-to-day world, and stimulated modes of thinking useful outside the classroom.

(3) *Relevant topics increase student interest and motivation.* Relevance which motivated students to intellectual activity came from the individual challenge to conquer a puzzle, from special interests, or from a student's identification with his career.

The more radical hypotheses involving free flowing discussion and eliminating grades proved unworkable, although less extreme policies did work. Observations regarding these hypotheses indicate that:

(1) *Pursuing student interests as they arise in class discussions yields scattered and frequently nonfunctional learnings.* The general pursuit of interests at the moment that they are expressed in class is similar to "impulse buying" at a market, which may result in expensive and useless purchases. However, a combination of student input with teacher guidance often worked well.

(2) *Course grades are, indeed, functional at these developing institutions.* Students used strategies to obtain good grades independent of the learning which the grades were intended to measure. A not unexpected result, however, was that students often responded with interest and application to learning tasks that were rewarded with good grades.

5. Conclusion. The T.C.C.P. has been effective in producing faculty innovation, in increasing student self-confidence, and in developing lively, thoughtful classes. Conclusions from the classroom experiments indicate that students can be drawn into thoughtful activities by puzzles and intellectual challenges, that students can enjoy college mathematics classes, and that an active, noisy classroom may be more productive than an orderly quiet one. The T.C.C.P. and similar programs at developing colleges need the understanding and support of the members and leaders of the M.A.A. Members of such professional organizations have an important influence on the priorities of the scientific community. Such programs benefit not only the developing institutions but all of American education.

PROBLEMS AND SOLUTIONS

EDITED BY EMORY P. STARKE

All problems (both elementary and advanced) proposed for inclusion in this Department should be sent to E. P. Starke, 1000 Kensington Ave., Plainfield, N. J. 07060. Proposers of problems are urged to enclose any solutions or information that will assist the editors. Ordinarily, problems in well-known textbooks and results in generally accessible sources are not appropriate for this Department. No solutions (except those accompanying proposals) should be sent to Professor Starke.

ELEMENTARY PROBLEMS

Solutions of Elementary Problems should be sent to Problems Group, Mathematics Department, University of Maine, Orono, ME 04473. To facilitate their consideration, solutions of Elementary Problems in this issue should be typed (with double spacing) and should be mailed before January 31, 1972. Contributors (in the United States) who desire acknowledgement of receipt of their solutions are asked to enclose self-addressed stamped postcards.

An asterisk (*) means neither the proposer nor the editors supplied a solution.

E 2313.* *Proposed by Sidney Heller, Brookhaven National Laboratory*

Show that

$$\sum_{i_m=1}^{n-m+1} \sum_{i_{m-1}=i_m+1}^{n-m+2} \cdots \sum_{i_3=i_4+1}^{n-2} \sum_{i_2=i_3+1}^{n-1} \sum_{i_1=i_2+1}^{n} 1 = \binom{n}{m}.$$

E 2314.* *Proposed by A. K. Austin, The University, Sheffield, England*

Prove or disprove that it is possible to find a convex polygon and three translations of it in the plane which form a Venn diagram for four sets (i.e., they form 16 connected regions and no three edges pass through the same point).

E 2315. *Proposed by Richard Stanley, Harvard University*

Let $f(n)$ be the number of ways an $(n+1)$-sided convex polygon can be divided into regions by diagonals not intersecting in the interior of the polygon. The trivial division, that is the division using no diagonals, is to be counted, so that $f(1)=1$, $f(2)=1$, $f(3)=3$, $f(4)=11$, etc. Find the generating function $F(x)=\sum f(n)x^n$, and find an asymptotic formula for $f(n)$.

E 2316. *Proposed by R. S. Luthar, University of Wisconsin at Janesville*

Show that

$$\phi(n^2) + \phi(n^2+2n+1) < 2n^2,$$

where n is any integer >2.

E 2317. *Proposed by R. S. Luthar, University of Wisconsin at Janesville*

Find all pairs of natural numbers m, n such that

$$\phi(mn) = \phi(m) + \phi(n).$$

E 2318. *Proposed by Thomas Hughes, Arlington, Texas*

Suppose that a machine is constructed to shuffle an ordinary 52-card deck in the same manner each time. How efficient could this machine be? That is, what is the maximum number of shuffles that could occur before the deck is returned to its original order?

SOLUTIONS OF ELEMENTARY PROBLEMS

The Integral of the Reciprocal of a Polynomial

E 2236 [1970, 522; 1971, 408]. *Proposed by M. S. Klamkin, Ford Scientific Laboratory, and D. J. Newman, Yeshiva University*

Show that if the integral of the reciprocal of a nonconstant polynomial is a rational function, then the polynomial must be of the form $(ax+b)^n$.

II. *Comment and solution by L. R. Abramson, Riverside Research Institute, New York.* The published solution I is in error: if f, g and p are polynomials such that f/g is in its lowest terms and $(f/g)'=1/p$, then f need not be constant, for it is not necessarily true that each of fg', gf' divides g^2. For example, let $f(x)=x-1$, $g(x)=x$, and $p(x)=1/x^2$.

The solution may be corrected as follows. Evidently $\deg f \leqq \deg g$. If $\deg f = \deg g$, then we can write $f/g=c+f_1/g$, where $\deg f_1<\deg g$. Since f_1/g is another antiderivative for p, there is no loss of generality in assuming that $\deg f<\deg g$. Let the leading terms of f and g be respectively ax^s and bx^t. Then the leading term of $gf'-fg'$ is $ab(s-t)x^{s+t-1}$, since $s-t\neq 0$. Inspection rules out the cases $s=0$, $t=1$ and $s=1$, $t=0$; hence $s+t\geqq 2$, and so $s+t-1\geqq 1$. As in the published solution every m-fold root of $gf'-fg'$ is an $(m+1)$-fold root of g. Thus $t=\deg g\geqq(s+t-1)+d$, where d is the number of distinct roots of $gf'-fg'$. But $d\geqq 1$, whence $s=0$ and $d=1$. In other words, f is constant and g' has exactly one distinct linear factor; i.e., $g(x)=(ax+b)^n$ for some $n\geqq 2$.

Hermitian Matrix

E 2254 [1970, 882]. *Proposed by Marvin Marcus, University of California, Santa Barbara*

Let H be an $n\times n$ hermitian matrix and let $|H|$ be the matrix obtained from H by replacing each entry by its absolute value. Show that if $H\geqq 0$ (i.e., H is positive semi-definite) and $n\leqq 3$, then $|H|\geqq 0$. Show that for each $n\geqq 4$ there exists an $H\geqq 0$ such that $|H|$ is indefinite.

Solution by the proposer. Let $H=(h_{ij})$ be an $n\times n$ hermitian matrix for which $H\geqq 0$. Let $|H|=(|h_{ij}|)$. The following well-known theorem about hermitian matrices will be used: If B is an $n\times n$ hermitian matrix, then $B\geqq 0$ if and only if the principal minors of B are all nonnegative. For $n=1$ the theorem makes our result obvious. For $n=2$ it is easy to show that $\det H=\det|H|$ since H is hermitian and $H\geqq 0$. For $n=3$ the theorem can be applied after it is shown that $\det|H|-\det H\geqq 0$. This follows since

$$\det|H| - \det H = 2[|h_{21}h_{32}h_{13}| - \operatorname{Re}(h_{21}h_{32}h_{13})] \geqq 0.$$

For $n=4$ the following matrix suffices

$$H_4 = \begin{bmatrix} 1 & \frac{1}{\sqrt{3}} & 0 & -\frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{3}} & 1 & \frac{1}{\sqrt{3}} & 0 \\ 0 & \frac{1}{\sqrt{3}} & 1 & \frac{1}{\sqrt{3}} \\ -\frac{1}{\sqrt{3}} & 0 & \frac{1}{\sqrt{3}} & 1 \end{bmatrix}.$$

For $n>4$ let $H=H_4\oplus I_{n-4}$.

Also solved by D. M. Bloom, R. M. Krause, and J. P. McLean.

A Continuous Function on R^n

E 2255 [1970, 882]. *Proposed by T. E. Mott, State University College, Buffalo, N.Y.*

Let $f(x_1, \cdots, x_n)$ be a real valued function on an open set $G\subseteq R^n$ and let $v_i=(\lambda_{i,1}, \cdots, \lambda_{i,n})$, $i=1, \cdots, n$, be linearly independent vectors in R^n. If the function f is continuous along that portion of every line passing through G and parallel to v_i, $i=1, \cdots, n$, and f is monotonic along each of these lines (the direction of monotonicity depending upon the choice of line), then $f(x_1, \cdots, x_n)$ is continuous in G.

Solution by David Spear, City College, New York. We assume without loss of generality that $\mathbf{v}_i$ is the ith coordinate axis and that f is monotonically increasing along each $\mathbf{v}_i$. We know that R^n can be partially-ordered as follows: if $\mathbf{a}=(a_1, \cdots, a_n)$ and if $\mathbf{b}=(b_1, \cdots, b_n)$, then $\mathbf{a}\geqq\mathbf{b}$ whenever $a_i\geqq b_i$ for $i=1, 2, \cdots, n$. The problem is now to show that if f is individually continuous in each of its variables and if $f(\mathbf{a})\geqq f(\mathbf{b})$ whenever $\mathbf{a}\geqq\mathbf{b}$, then f is continuous.

Let $\mathbf{x}_0$ be a point of G. We shall show that f is continuous at $\mathbf{x}_0$. We make the further simplifying assumptions that $\mathbf{x}_0=0$ and that $f(\mathbf{x}_0)=0$.

Since G is open, we can find $h>0$ so that the box

$$\{(x_1, \cdots, x_n): |x_i| < h \quad \text{for } i = 1, \cdots, n\}$$

is completely contained within G. Let $\epsilon>0$ be given. By assumption, there exists p_1, $0<p_1<h$, such that $0 \leqq f(p_1, 0, 0, \cdots, 0) < \epsilon/n$. Having found p_1, we can find p_2, $0<p_2<h$, such that $0 \leqq f(p_1, p_2, 0, \cdots, 0) - f(p_1, 0, 0, \cdots, 0) < \epsilon/n$ and hence $0 \leqq f(p_1, p_2, 0, \cdots, 0) < 2\epsilon/n$. Proceeding inductively we can find $\mathbf{p} = (p_1, p_2, \cdots, p_n)$ such that $0<p_i<h$ for $i=1, 2, \cdots, n$ and such that $0 \leqq f(\mathbf{p}) < n\epsilon/n = \epsilon$.

In the same way we can find $\mathbf{q} = (q_1, q_2, \cdots, q_n)$ such that $-h<q_i<0$ for $i=1, \cdots, n$ and such that $0 \geqq f(\mathbf{q}) > -\epsilon$. Now let B be the open box

$$B = \{(x_1, \cdots, x_n): q_i < x_i < p_i \quad \text{for } i = 1, \cdots, n\}.$$

If $\mathbf{x} \in B$, then $\mathbf{q} \leqq \mathbf{x} \leqq \mathbf{p}$ and so we have $-\epsilon < f(\mathbf{q}) \leqq f(\mathbf{x}) \leqq f(\mathbf{p}) < \epsilon$ implying that f is continuous at $\mathbf{0}$.

Also solved by J. B. Wilker, and by the proposer (who refers to R. L. Kruse and J. J. Deely, *Joint continuity of monotonic functions*, this MONTHLY, 74 (1969) 74–76).

A Small Function with Large Gradient

E 2256 [1970, 882]. *Proposed by H. Kestelman, University College, London, England*

Determine a real function f on R^n so that $|f(x)| \leqq \|x\|$ for all x, where $\|x\| = (x_1^2 + \cdots + x_n^2)^{1/2}$ and $\|\operatorname{grad} f(0)\| = n^{1/2}$.

I. *Solution by Harry Lass, California Institute of Technology.* Let $f(0) = 0$, and for $x \neq 0$ let

$$f(x) = \sum_{i=1}^{n} x_i^3/\|x\|^2.$$

II. *Solution by Rhodes Peele, University of North Carolina.* Let

$$f(x) = \begin{cases} \|x\| & \text{if } \sum_{i=1}^{n} x_i \geqq 0, \\ -\|x\| & \text{if } \sum_{i=1}^{n} x_i < 0. \end{cases}$$

III. *Solution by E. F. Schmeichel, Itasca, Illinois.* Let f vanish except on the coordinate axes and let $f(x) = x_i$ if x is on the ith coordinate axis.

Solution I was also submitted by the proposer. Solution III was also submitted by G. A. Heuer, Barbara Keller, and J. B. Wilker.

Keller and Peele observe that unless $n=1$ it is not possible to have a solution which is differentiable at the origin.

A Harmonic Identity

E 2257 [1970, 883]. *Proposed by I. Kaucký, Bratislava, Czechoslovakia*

Prove the identity

$$\sum_{k=1}^{2n-1}(-1)^{k-1}\binom{2n-1}{k}^{-1}\sum_{j=1}^{k}\frac{1}{j}=\frac{2n}{2n+1}\sum_{k=1}^{2n}\frac{1}{k}\,.$$

Solution by M. G. Greening, University of New South Wales, Australia. We note that

$$\binom{2n-1}{k}^{-1}=\frac{2n}{2n+1}\left[\binom{2n}{k}^{-1}+\binom{2n}{k+1}^{-1}\right],$$

so that the given identity is equivalent to the following:

$$\text{(1)}\qquad \sum_{k=1}^{2n-1}(-1)^{k-1}\left[\binom{2n}{k}^{-1}+\binom{2n}{k+1}^{-1}\right]\sum_{j=1}^{k}\frac{1}{j}=\sum_{k=1}^{2n}\frac{1}{k}\,.$$

Consider the left-hand side of (1) above. Break it into two sums; in the first of these, split off the first term and in the second, split off the last (i.e., the $(2n-1)$st) term and change the summation index, letting $i=k+1$. This gives the following:

$$\text{(2)}\qquad \begin{aligned}&(2n)^{-1}+\sum_{k=2}^{2n-1}(-1)^{k-1}\binom{2n}{k}^{-1}\sum_{j=1}^{k}\frac{1}{j}+\sum_{i=2}^{2n-1}(-1)^{i}\binom{2n}{i}^{-1}\sum_{j=1}^{i-1}\frac{1}{j}\\&+(-1)^{2n-2}\sum_{j=1}^{2n-1}\frac{1}{j}\,.\end{aligned}$$

We can combine the first and fourth terms of (2) to give simply $\sum_{j=1}^{2n}1/j$; the second and third terms also can be combined and (2) becomes

$$\text{(3)}\qquad \sum_{k=2}^{2n-1}(-1)^{k-1}\binom{2n}{k}^{-1}\frac{1}{k}+\sum_{j=1}^{2n}\frac{1}{j}\,.$$

But if we examine the first term of (3) we see that

$$\sum_{k=2}^{2n-1}(-1)^{k-1}\binom{2n}{k}^{-1}\frac{1}{k}=\frac{1}{2n}\sum_{k=2}^{2n-1}(-1)^{k-1}\binom{2n-1}{k-1}^{-1},$$

which is zero because of the fact that the term of the sum for $k=k_0$ is cancelled by the term for $k=2n+1-k_0$. This establishes the identity.

Also solved by L. Carlitz, R. M. Krause, Harry Lass, C. L. Sabharwal, B. L. R. Shawyer, J. R. Ventura, Jr., and the proposer.

Editorial Comment. The present formula should be compared with Formula 2.18 of H. W. Gould, *Combinatorial Identities* (Morgantown, W. Va., 1959) which asserts

$$\sum_{k=1}^{2n-1}(-1)^{k-1}\binom{2n}{k}^{-1}\sum_{j=1}^{k}\frac{1}{j}=\frac{n}{2(n+1)^2}+\frac{1}{2(n+1)}\sum_{k=1}^{2n}\frac{1}{k}\,.$$

A Number-theoretic Summation

E 2258 [1970, 883]. *Proposed by Arthur Marshall, Madison, Wisc.*

For each natural number k, let N_k be the kth number in the sequence which consists only of primes and products of consecutive primes, taken in natural order. Let the Möbius function be defined as usual: $\mu(1)=1$, and for $n>1$, $\mu(n)=(-1)^r$, where r is the number of primes dividing (squarefree) n. Prove

$$\sum_{k=1}^{\infty}\frac{\mu(N_k)}{N_k}=-\infty.$$

Solution by Allen Stenger, Student, Emory University. For each prime p, let $p, q, r, \cdots$ be successive primes. Note that, given p, the rest are uniquely determined. Since $p<q<r\cdots$, we have also $1/pq<1/p^2$, $1/pqr<1/p^3$, etc. Hence we get

$$\begin{aligned}\sum_{k=1}^{j}\frac{\mu(N_k)}{N_k}&=-\sum_{p\le N_j}\frac{1}{p}+\sum_{pq\le N_j}\frac{1}{pq}-\sum_{pqr\le N_j}\frac{1}{pqr}+\cdots\\&<-\sum_{p\le N_j}\frac{1}{p}+\sum_{pq\le N_j}\frac{1}{p^2}+\sum_{pqr\le N_j}\frac{1}{p^3}+\cdots\\&<-\sum_{p\le N\cdot}\frac{1}{p}+\sum_{p}\frac{1}{p^2}+\sum_{p}\frac{1}{p^3}+\cdots\\&<-\sum_{p\le N_j}\frac{1}{p}+\sum_{p}\frac{1}{p(p-1)}<-\sum_{p\le N_j}\frac{1}{p}+1.\end{aligned}$$

Now let $j\to\infty$ so that $N_j\to\infty$ and hence the right-hand side of the above approaches $-\infty$ since $\sum 1/p$ diverges.

Also solved by Anders Bager (Denmark), Joseph Gillis (Israel), M. G. Greening (Australia), Emil Grosswald, Simeon Reich (Israel), E. F. Schmeichel, David Spear, H. H. Thoyre, Charles Wexler, P. H. Young, and the proposer.

Commuting Powers in a Group

E 2259 [1970, 1007]. *Proposed by J. R. Isbell, State University of New York at Buffalo*

A group in which all uth powers commute with each other and all vth powers commute with each other, u and v relatively prime, is abelian.

Solution by B. M. Green, Monmouth, Oregon. Let x and y be integers such that $xu+yv=1$ and let a and b be arbitrary elements of the group. Then

$$(a^ub^v)^{xu}=a^u(b^va^u)^{(xu-1)}b^v=(b^va^u)^{xu}.$$

Similarly $(a^ub^v)^{yv}=(b^va^u)^{yv}$. Thus

$$a^ub^v=(a^ub^v)^{(xu+yv)}=(b^va^u)^{(xu+yv)}=b^va^u,$$

so that all uth powers commute with all vth powers. Therefore

$$ab = a^{(xu+yv)}b^{(xu+yv)} = ba.$$

Also solved by Ram Awtar (India), G. W. Fehlhaber, Ralph Garfield, M. G. Greening (Australia), C. J. Leska, D. E. Manes, R. A. Moore, Problems Seminar of Luther College, H. D. Ruderman, Wolfe Snow, R. Z. Vause (Saudi Arabia), J. H. Webb (South Africa), E. T. Wong, and the proposer.

Editorial Note. Ruderman gives a generalization to the case where u and v need not be relatively prime. He shows that if $d=(u, v)$, then $(a^d b^d)^d = (b^d a^d)^d$ for all a and b.

A Transcendental Limit

E 2260 [1970, 1007]. *Proposed by Marlow Sholander, Case Western Reserve University*

It is given for every real $n>0$ that three times the area between $y=1$ and $y=x^n$, $0\leqq x\leqq 1$, equals the area between $y=1$ and $y=x^n$, $1\leqq x\leqq a_n$. Find $\lim_{n\to 0} a_n$.

Solution by Eddy Smet, Graduate Student, University of Western Ontario. For clarity, we replace the variable n by the variable t. Evaluation of the integrals shows that a_t must satisfy

$$a_t\left[\frac{a_t^t - (t+1)}{t}\right] = 2.$$

For $t>0$ and $x\geqq 1$, define the function f_t by

$$f_t(x) = x\left[\frac{x^t - (t+1)}{t}\right].$$

It is easy to see that for each t, f_t increases (strictly) from -1 to ∞, so that a_t is well-defined as the unique solution to the equation $f_t(x)=2$.

For every $x\geqq 1$ define $f(x)$ by $f(x)=\lim_{t\to 0} f_t(x)$. Standard methods show that $f(x)=x(\log x-1)$. Evidently f also increases (strictly) from -1 to ∞, so that we can let a be the unique solution to the transcendental equation $f(x)=2$. [*Editorial note*: Numerical methods give $a=4.3191365663\cdots$.] We shall show that $a_t\to a$ as $t\to 0$.

Suppose that lim inf $a_t<a$. Then there exists $M<a$ and a sequence $\{t_k\}$ of positive numbers such that $t_k\to 0$ and such that $a_{t_k}\leqq M$ for all k. Then $2=f_{t_k}(a_{t_k}) \leqq f_{t_k}(M)$ so that $2\leqq \lim_k f_{t_k}(M)=f(M)<f(a)=2$, a contradiction. Thus lim inf $a_t \geqq a$. Similarly we can establish that lim sup $a_t\leqq a$, so that lim inf $a_t=$ lim sup $a_t =a$, and thus $a_t\to a$.

Also solved by K. F. Andersen, G. W. Berg & R. L. Young, Frederick Carty, R. N. Chilcote, Jordi Dou (Spain), John Flaig, R. G. Griswold, Emil Grosswald, Ellen Hertz, M. Hirschhorn (Scotland), Harry Lass & Peter Gottlieb, C. S. Ogilvy, Simeon Reich (Israel), H. D. Ruderman, St. Olaf College Students, Wolfe Snow, Robert Vinyard, P. H. Young, and the proposer.

Editorial Comment. A great many incorrect solutions to this problem were received. The most common mistake was the assumption, without any justification, that lim a_t exists. A surprising number of solvers evaluated the wrong limit.

Decreasing Likelihood

E 2261 [1970, 1007]. *Proposed by R. M. Meyer, SUNY College at Fredonia, N.Y.*

Show that for each integer m, $0 \leqq m < n$, the function

$$L_m(p) = \sum_{j=0}^{m} \binom{n}{j} p^j (1-p)^{n-j}$$

is strictly decreasing for $0 < p < 1$.

Solution by S. M. Rohde, GM Research Laboratories, Warren, Mich. Evidently $1 - L_m(p)$ is the probability of getting more than m successes out of n Bernoulli trials with probability of success p; clearly then $1 - L_m(p)$ is a strictly increasing function of p, so that $L_m(p)$ is a strictly decreasing function of p.

Also solved by Marcia Ascher, D. J. Bordelon, W. D. Bouwsma, L. Carlitz, Frederick Carty, Peter Ellis, Peter Enis, Michael Goldberg, Peter Gottlieb & Henry Lass, M. G. Greening (Australia), Emil Grosswald, Robert Heller, Ellen Hertz, J. C. Hickman, M. Hirschhorn (Scotland), W. T. Hodson, A. R. Jiminez, Elgin Johnston, P. M. Kannan, N. J. Kuenzi, L. Kuipers, Joel Levy, R. B. Lind, Luther College Problems Seminar, D. E. Manes, J. V. Michalowicz, D. L. Muench, Oscar Ocelot, A. J. Patsche, M. E. Price, J. G. Rau, Simeon Reich (Israel), G. S. Rogers, D. S. Rubin, E. M. Scheuer, E. F. Schmeichel, F. G. Schmitt, Jr., Kim Scorp, H. T. Sedinger, B. L. R. Shawyer, Sid Spital, St. Olaf College Students, Jim Tattersall, M. W. Varano, J. R. Ventura, Jr., Julius Vogel, M. Waterman, Charles Wexler, P. H. Young, and the proposer.

Editorial Comment. A number of solvers note that

$$L_m(p) = (n-m) \binom{n}{m} \int_0^{1-p} x^{n-m-1}(1-x)^m dx,$$

which is strictly decreasing as a function of p for $0 < p < 1$ since the integrand is positive. Another popular method is straight-forward differentiation of $L_m(p)$ to show that

$$\frac{d}{dp}(L_m(p)) = -n \binom{n-1}{m} p^m (1-p)^{n-m-1} < 0 \qquad \text{for } 0 \leqq m < n.$$

ADVANCED PROBLEMS

All solutions of Advanced Problems should be sent to J. Barlaz, Rutgers—The State University, New Brunswick, N. J. 08903. Solutions of Advanced Problems in this issue should be typed (with double spacing) on separate, signed sheets and should be mailed before January 31, 1972. Contributors (in the United States) who desire acknowledgement of receipt of their solutions are asked to enclose self-addressed, stamped postcards.

An asterisk (*) means neither the proposer nor the editors supplied a solution.

5814. *Proposed by A. C. Segal, University of Alabama, Birmingham*

It is known that a subset of the real line which is either first category or measure zero must have a void intersection with uncountably many Lebesgue cosets (i.e., cosets modulo the subgroup of rationals). Prove or disprove the converse.

5815. *Proposed by L. W. Shapiro, Howard University*

Show there are no groups G of order p^{2n} with center and commutator subgroup of order p, where p is a prime.

5816. *Proposed by Solomon Leader, Rutgers—The State University*

Let P be a nonempty, finite set with p members, and Q be a finite set with q members. Let $N_k(p, q)$ be the number of binary relations of cardinality k with domain P and range Q. (Equivalently, $N_k(p, q)$ is the number of $p\times q$ matrices of 0's and 1's with exactly k entries equal to 1 and no row or column identically 0.) Compute

$$\sum_{k=1}^{pq} (-1)^{k-1} N_k(p, q).$$

5817. *Proposed by M. F. Neuts, Purdue University*

Let $f(t)$ be the characteristic function of a probability distribution $F(\cdot)$ whose $(n+1)$st moment is finite. For all $\lambda>0$ the integral $I(\lambda)=\int_0^\infty e^{-\lambda t}f(t)dt$ exists. Prove that

$$\lim_{\lambda\to+\infty} \frac{1}{I(\lambda)} \sum_{\nu=0}^{n} i^\nu \lambda^{-\nu-1}\mu_\nu' = 1,$$

where μ_ν' is the ν-th moment of $F(\cdot)$.

As a particular case, obtain the classical asymptotic expansion

$$1-\Phi(\lambda) \sim \frac{1}{\sqrt{2\pi}} e^{-\lambda^2/2} \sum_{\nu=1}^{n} (-1)^\nu \lambda^{-2\nu-1}(1\cdot 3 \cdot\cdot\cdot (2\nu-1))$$

for the normal distribution function, as $\lambda\to+\infty$.

5818. *Proposed by Erwin Just, Bronx Community College*

Let $q\geqq 5$ be an integer of one of the forms $6n\pm 1$. Must there exist a prime p and an integer x for which

$$x^{q-1}-x+1\equiv 0 \pmod{p} \quad\text{and}\quad x^q\equiv 1 \pmod{p}?$$

5819*. *Proposed by S. Abhyankar, Purdue University*

Let $f(x)$ and $g(x)$ be nonconstant polynomials in one variable x with coefficients in a field K. Let m be the degree of $f(x)$ and n the degree of $g(x)$, where $1<m<n$. Assume that x can be expressed as a polynomial in $f(x)$ and $g(x)$, i.e., there exists a polynomial $h(y, z)$ in two variables y and z with coefficients in K such that $h(f(x), g(x))=x$.

(1) Show that m and n must have a common factor greater than 1.
(2) Show that $(m, n)\neq(4, 6)$.
(3) Is it necessarily true that $(m, n)\neq(4, 10)$?
(4) More generally, is it true that m must divide n?

SOLUTIONS OF ADVANCED PROBLEMS

Curvatures at Intersections of Conics, and of Quadric Hypersurfaces

5745 [1970, 656]. *Proposed by Daniel Pedoe, University of Minnesota*

Q is a nonspecialized quadric hypersurface in Euclidean space of n dimensions, L a hyperplane which intersects Q in a nonspecialized quadric of one less dimension. We consider the pencil of quadric hypersurfaces $Q+kL^2=0$, and suppose that Q_1, Q_2, Q_3, Q_4 are any four of them; T_1, T_2, T_3, T_4 the total Gaussian curvatures at any point of the quadric of contact $Q=L=0$; M_1, M_2, M_3, M_4 the mean Gaussian curvatures of the respective quadrics at the same point. Prove that the two cross-ratios

$$\{T_1, T_2, T_3, T_4\}, \qquad \{M_1, M_2, M_3, M_4\}$$

are equal. Deduce that if two conics touch at two points P_1, P_2, then the ratio of their curvatures is the same at both points.

Solution by the proposer. The key to the solution is to see that each of the two given cross-ratios is equal to the cross-ratio of the values of k which give the four quadrics. Details are given in D. Pedoe, *A remark on a property of a special pencil of quadrics*, Proc. Camb. Phil. Soc., 38(1942) 235. Since the cross-ratio of the values of k is unaltered by taking a section of the configuration by a hyperplane, a similar equality is obtained for the curvatures in a space of lower dimension, until finally we obtain this result: If Q_1, Q_2, Q_3, Q_4 are four conics of a pencil which touch each other at the two points P_1, P_2, then the cross-ratio of the curvatures of the four conics at P_1 is equal to the cross-ratio of the curvatures of the four conics at P_2. The deduction suggested in the problem follows from the reasonable assumption that the total and mean curvatures of the degenerate quadric $L^2=0$ at any point on its section with $Q=0$ are both infinite, and that a tangent hyperplane to $Q=0$, considered as part of a reducible quadric, has both total and mean curvature equal to zero. The deduction is due to B. Segre, Rendiconti Accad. Lincei, 9(1929) 970–974.

Approximation of Bounded Monotonic Function

5750 [1970, 775]. *Proposed by John Horvath, University of Maryland*

Let f be a decreasing, bounded function, defined on the interval $0 \leqq x \leqq 1$ of the real line. Prove that there exists a sequence (f_n) of continuous, decreasing functions having the same bounds as f, which converges almost everywhere to f.

Solution by W. C. Waterhouse, Cornell University. Since f is monotone it has only countably many points of discontinuity, say $d_1, d_2, \cdots$. For each n partition $[0, 1]$ using the points $0, 1/n, 2/n, \cdots, 1$ and $d_1, d_2, \cdots, d_n$; and let f_n be the function which is linear on the resulting subintervals and agrees with f

at the partition points. Clearly f_n is a continuous decreasing function having the same bounds as f; and f_n actually converges to f *everywhere.*

Indeed, we have $f_n(d_i)=f(d_i)$ for all $n\geqq i$. If c is a point of continuity, and $\epsilon>0$, choose $\delta>0$ with $|f(x)-f(c)|<\epsilon/2$ for $|x-c|<\delta$. Then as soon as $n>1/\delta$, the subinterval that contains c will lie within $(c-\delta, c+\delta)$, and hence $|f_n(c)-f(c)|<\epsilon$.

Also solved by Linda W. Brinn, R. A. Christiansen, L. E. Clarke (England), D. E. Daykin (England), M. A. B. Deakin (New Guinea), C. R. Diminnie, D. Ž. Djoković, Robert Fefferman, D. A. Hejhal, A. A. Jagers (Netherlands), Eleanor G. Jones, G. C. T. Kung, Joel Levy, E. A. Memmott, Ka Menehune, F. P. Miller, Jr., K. R. Milliken, Nicholas Passell, Henry Ricardo, E. F. Schmeichel, P. W. Smith, J. R. Trollope, R. M. Warten, Mark Yu, and the proposer.

Editorial Notes. (1) Hejhal, Memmott and Ricardo show that f_n may be chosen absolutely continuous by using integral means of f over $(x, x+1/n)$. Memmott also shows that the condition of monotonicity on f may be relaxed if we do not require f_n to be monotonic; e.g., f summable, f_n converging almost everywhere to f, in particular at all points where f is continuous.

(2) Using a method which applies also to an analogous problem on page 197 of W. Rudin, *Complex and Real Analysis*, Diminnie shows that f_n may be chosen infinitely differentiable in $[0, 1]$.

A Combinatoric Evaluation

5752 [1970, 775]. *Proposed by Kesiraju Satyanarayana, Rajahmundry, India*

If

$$\sigma_{m,r}=\sum_{p=0}^{m}(-1)^p\binom{2m+1}{p}^r\cdot\sum_{x=p+1}^{2m+1-p}\frac{1}{x}$$

is expressed as a fraction in lowest terms, show that

(1) $\sigma_{m,1}=1/(2m+1)$.

(2) For $m=1$ to 7, the following results hold:

(i) $\sigma_{m,2}$, $\sigma_{m,3}$ have the sign of $(-1)^m$.

(ii) The numerator of $|\sigma_{m,2}|$ is a power of 2.

(iii) The numerator of $|\sigma_{m,3}|/(2m+1)!$ is the product of some successive prime numbers immediately following $2m+1$.

(3) Do the results in (2) hold in general?

I. *Partial solution by the proposer.* (1) The given equation for $r=1$ may be written as

$$\sigma_{m,1}=\sum_{r=1}^{m}\left[\frac{1}{r}+\frac{1}{2m+2-r}\right]\left[\sum_{s=0}^{r-1}(-1)^s\binom{2m+1}{s}\right]$$
$$+\frac{1}{m+1}\left[\sum_{s=0}^{m}(-1)^s\binom{2m+1}{s}\right].$$

Using the identity $(1-x)^n\cdot(1-x)^{-1}=(1-x)^{n-1}$, we obtain

$$\sum_{s=0}^{t}(-1)^s\binom{n}{s}=(-1)^t\binom{n-1}{t}.$$

Also

$$\sum_{r=1}^{m}(-1)^{r-1}\left[\frac{1}{r}+\frac{1}{2m+2-r}\right]\binom{2m}{r-1}$$

$$=\sum_{r=1}^{m}\frac{1}{2m+1}\left[(-1)^{r-1}\binom{2m+1}{r}+(-1)^{2m+1-r}\binom{2m+1}{2m+2-r}\right]$$

$$=\frac{1}{2m+1}\left[\sum_{r=1}^{m}(-1)^{r-1}\binom{2m+1}{r}+\sum_{t=2m+1}^{m+2}(-1)^{t-1}\binom{2m+1}{t}\right],$$

where $t=2m+2-r$. We have

$$\frac{1}{m+1}\sum_{s=0}^{m}(-1)^{s}\binom{2m+1}{s}=\frac{(-1)^m}{m+1}\cdot\frac{(2m)!}{m!m!}=\frac{(-1)^m}{2m+1}\binom{2m+1}{m+1},$$

and therefore

$$\sigma_{m,1}=\frac{1}{2m+1}\left[\sum_{r=1}^{2m+1}(-1)^{r-1}\binom{2m+1}{r}\right]=\frac{1}{2m+1}\left\{1-\sum_{r=0}^{2m+1}(-1)^{r}\binom{2m+1}{r}\right\}$$

$$=1/(2m+1).$$

(2) The assertions in (2) may be verified by direct calculations. The number of primes in the numerator of $|\sigma_{m,3}|/(2m+1)!$ is $(m+1)$ except for $m=6$ in which case it is m.

II. *Addendum by Leonard Carlitz, Duke University.* With

$$\sigma_k=1+\frac{1}{2}+\frac{1}{3}+\cdots+\frac{1}{k}=\int_0^1\frac{1-x^k}{1-x}\,dx,\qquad (k\geqq 1),\quad \sigma_0=0,$$

the following formulas can be obtained for $\sigma_{m,r}$:

$$\text{(a)}\quad \sigma_{m,r}=-\sum_{k=0}^{2m+1}(-1)^k\binom{2m+1}{k}^r\sigma_k,$$

$$\text{(b)}\quad \sigma_{m,r}=-\sum_{k=0}^{2m+1}(-1)^k\binom{2m+1}{k}^r\int_0^1\frac{1-x^k}{1-x}\,dx,$$

$$\text{(c)}\quad \sigma_{m,2}=(-1)^m\frac{2^{4m}}{(2m+1)\binom{2m}{m}},$$

$$\text{(d)}\quad \sigma_{m,3}=\sum_{j=0}^{m}(-1)^j\frac{(2m+j+1)!}{j!j!(2m-j+1)!}\cdot\frac{1}{2m-2j+1}.$$

and from this last it follows that the sign of $\sigma_{m,3}$ is $(-1)^m$.

These become $\bar{x}$ and x' respectively when adjustments are made which are suggested by the exact values of x_n for $n=3$ and 4, namely 2 and $\sqrt{1+\sqrt{2}}$ respectively. Some comparative values are:

n	x_n	$\bar{x}$	x'
3	2.00000	2.00713	1.99590
4	1.55377	1.55597	1.55318
5	1.38481	1.38558	1.38466
6	1.29524	1.29546	1.29519
7	1.23962	1.23959	1.23959
8	1.20168	1.20153	1.20166
9	1.17413	1.17393	1.17411
10	1.15322	1.15298	1.15320

Also solved by Edgar Karst, E. M. Stone, Takashi Tamura (Japan), L. E. Vogler and Philip H. Young.

The Perimeter of an Ellipse

5754 [1970, 890; 1971, 202]. *Proposed by M. S. Klamkin, Ford Scientific Laboratory*

Let $L(a, c)$ equal the perimeter of an ellipse with semi-axes a and c with $a \geqq c$. Show that, if $a \geqq b$, then

$$L^2(a, c) - 16a^2 \geqq L^2(b, c) - 16b^2.$$

Solution by K. F. Andersen, Royal Roads Military College, Victoria, B.C. Parameterize the ellipse in the usual trigonometric manner and then let

$$u = u(t) = (a^2 \sin^2 t + c^2 \cos^2 t)^{1/2}, \qquad v = v(t) = (b^2 \sin^2 t + c^2 \cos^2 t)^{1/2}.$$

Then $u \geqq v$ since $a \geqq b$, and by Schwartz' inequality we obtain:

$$\begin{aligned} L^2(a, c) - L^2(b, c) &= (L(a, c) - L(b, c))(L(a, c) + L(b, c)) \\ &= \left(\int_0^{2\pi} (u - v)dt\right)\left(\int_0^{2\pi} (u + v)dt\right) \\ &= \left(\int_0^{2\pi} \frac{(u^2 - v^2)}{(u + v)} dt\right)\left(\int_0^{2\pi} \frac{(u^2 - v^2)}{(u - v)} dt\right) \\ &\geqq \left(\int_0^{2\pi} \frac{(u^2 - v^2)^{1/2}}{(u + v)^{1/2}} \cdot \frac{(u^2 - v^2)^{1/2}}{(u - v)^{1/2}} dt\right)^2 \\ &= \left(\int_0^{2\pi} (u^2 - v^2)^{1/2} dt\right)^2 \\ &= (a^2 - b^2)\left(2\int_0^{\pi} \sin t \, dt\right)^2 \\ &= 16(a^2 - b^2). \end{aligned}$$

Since equality holds in the Schwartz' inequality if and only if there are constants m, n, not both zero, such that $m(u-v)=n(u+v)$ almost everywhere, we have equality above if and only if $a=b$.

Also solved by E. D. Bolker, D. Borwein & B. L. R. Shawyer & B. Thorps, R. S. Castroll (Israel), J. H. E. Cohn (England), Josef Danes (Czechoslovakia), M. A. B. Deakin (New Guinea), Leon Gerber, Michael Goldberg, J. R. Hatcher, D. A. Hejhal, A. A. Jagers (Netherlands), R. B. Kirchner, Viktors Linis (Denmark), G. J. McRae (Australia), R. K. Meany, W. W. Meyer, R. H. C. Newton (Wales-United Kingdom), S. N. Rao, David Shelupsky, Michael Shimshoni, Marlow Sholander, J. E. Wilkins, Jr. and the proposer.

Editorial Note. The error in the first printing of the problem was noted by several solvers who derived interesting consequences from the false proposal. In particular, Bolker shows that the original statement implies $\pi^2=8$.

Limits of Operators on H

5755 [1970, 890]. *Proposed by G. G. Kiziak, University of Toronto*

It is well known that if a sequence of commuting normal operators T_n on a Hilbert space H converges strongly to an operator T, then T is normal. Show that the assumption that the operators commute cannot be dropped, even if the T_n are unitary. (This contradicts an assertion in Nagy and Forias, *Analyse Harmonique des Opêrateurs de l'espace de Hilbert*, p. 107, paragraph 2.)

I. *Solution by B. E. Cain, Iowa State University.* For each integer $n>0$ the operator T_n on the Hilbert space of square summable sequences, given by

$$T_n(x_1, x_2, \cdots) = (x_n, x_1, x_2, \cdots, x_{n-1}, x_{n+1}, x_{n+2}, \cdots),$$

is unitary because $(T_nx, T_ny)=(x, y)$ and T_n is invertible. The unilateral shift $S: x=(x_1, x_2, \cdots)\to(0, x_1, x_2, \cdots)$ is the strong limit of the T_n because

$$\|(T_n - S)x\| = \left[| x_n |^2 + \sum_{k=n+1}^{\infty} | x_k - x_{k-1} |^2 \right]^{1/2} \to 0,$$

but S is clearly not unitary.

II. *Comment by Bêla Sz.-Nagy, University of Szeged, Hungary.* It is known that every contraction operator T on an infinite dimensional Hilbert space H is the weak limit of a sequence of unitary operators $U_n (n=1, 2, \cdots)$ on H; see P. R. Halmos, *Normal dilations and extensions of operators*, Summa Brasil. Math., 2 (1950), 125–134, and B. Sz.-Nagy, *Suites faiblement convergentes de transformations normales de l'espace hilbertien*, Acta Math. Acad. Sci. Hung., 8 (1957), 295–302. If T is an isometry, we have $\|U_nx\|=\|x\|=\|Tx\|$, and hence weak convergence $U_n\rightharpoonup T$ implies strong convergence $U_n\to T$.

Also solved by S. L. Campbell, J. A. Goldstein, W. F. Moss, David Promislow, and the proposer.

REVIEWS

EDITED BY J. ARTHUR SEEBACH, JR., AND LYNN A. STEEN

with the assistance of the mathematics departments of St. Olaf and Carleton Colleges.

COLLABORATING EDITOR FOR FILMS: SEYMOUR SCHUSTER

Printed materials for review should be sent to: Book Review Editor, American Mathematical Monthly, St. Olaf College, Northfield, MN 55057. Films and correspondence relating to films should be sent to Seymour Schuster, Carleton College, Northfield, MN 55057.
*All unsigned material is written by one of the editors. A boldface capital **C** in the margin indicates that a review is based in part on classroom use. Professors willing to write such a review should inform the editor in order to avoid duplication.*

C *Optimization in Integers and Related Extremal Problems.* By Thomas L. Saaty. McGraw-Hill, New York, 1970. xv+295 pp. $16.50 (Telegraphic Review, October 1970.)

This book, the first of its kind, fills a definite need as a text or reference in the area of discrete optimization. The author's objective is to stimulate and excite the reader's interest and present some of the current ideas, techniques, and algorithms of optimization in integers, in both geometric and algebraic settings.

This reviewer found the chapters on geometric optimization and elementary applications amusing and within the reach of students at the early undergraduate level. Some of the material on integer programming, pseudo-boolean methods, and algorithms is more technical. The book abounds with motivation, applications, illustrations, exercises, and personal reflections. The author has limited the number of proofs to what he felt to be "adequate for teaching and for giving the student an idea of the kinds of proofs which occur in this general field." The format is pleasant; the style is lucid, witty, and quite informal.

The book is suitable for various types of courses in discrete and geometric optimization (parts of the book have been used in a short course on optimization at U.C.L.A.). A novel use of this book might be for an undergraduate seminar with active participation by the students.

A student studying this book carefully will encounter many ideas, learn interesting and imaginative techniques, and be greatly inspired to construct proofs himself and look up some of the cited references for more details. As a by-product, he will also learn something about model formulation, an aspect of mathematical education which has often been neglected in undergraduate curricula. Above all, he will certainly have a lot of fun, and learn how to cut a cake fairly at his next (optimally packed) party!

M. Z. NASHED, University of Wisconsin

When applying mathematics to problems in such fields as geometry and economics, one often encounters the need to optimize some function, subject to the condition that the solution be an integer or a set of integers. The aim of this text, intended for undergraduates with a background in advanced calculus and

linear algebra, is to present an introduction to a wide range of such problems. The book uses both algebraic and geometric methods. Despite the large quantity of interesting mathematics included here (there are chapters on classical methods, geometric optimization, Diophantine equations, and integer programming) the reviewer feels serious doubt about the effectiveness of this book as a text.

The most disturbing feature of the book is the arbitrary manner of organization and presentation, and the failure to provide information which seems appropriate. For example, the Fibonacci numbers are defined on page 10 and mentioned again briefly on page 29 in connection with a technique called "search by golden section." No reason is given the student for taking an interest in Fibonacci numbers, nor is any further information forthcoming about the golden section. Surely the latter deserves further mention when ninety pages of the book are devoted to geometric optimization. A second example of this sort of thing is the statement on page 164 that "the question of existence of solutions in positive integers to Fermat's famous equation, $x^n+y^n=z^n$ for $n>2$, is still not completely settled." Surely either too little or too much has just been said.

More serious than these matters is the question of the choice of theorems presented. Of course the author must decide on his approach to the material, but there seems to be an unusual number of arbitrary decisions made here. For example, about six pages of the chapter on geometric optimization is devoted to packing and filling in n-space, while the proof of Euler's formula is relegated to an exercise. Moreover, the beautiful way in which Euler's formula leads to a proof that there are at most five regular convex polyhedra in three dimensional space (see Courant and Robbins, *What is Mathematics?*) is not even mentioned, although this result has been given as an exercise earlier. Again, while there are more than twenty-five pages devoted to graphs and networks, the simple and instructive proof of the max-flow min-cut theorem (found for example in Hadley's *Linear Programming*) is omitted.

Another annoying feature of the book is the lack of clarity about the prerequisites required of the reader. Despite the assumed background of the reader, noted above, we find the following examples of seemingly divergent assumptions. On pages 21–22 the reader is introduced to metric topology, relative topology, subbases, bases, and the order topology of a partially ordered set. On page 89 he is suddenly expected to be familiar with quantifiers. On page 202 he is told that there are competitive problems which cannot be treated by game theory, a statement whose relevance is not clear since game theory is never mentioned elsewhere in the book. Finally, on page 234 the reader is informed that the purpose of the exposition of Gomory's algorithm for integer programming is "to open the door slightly · · · so the student · · · will approach the subject with a modicum of sophistication." One feels that the student's sophistication is probably sufficient for him to find this insulting.

We do not wish to suggest that this book is without merit. One can learn much interesting mathematics from it. For example, the brief third chapter,

"Some Elementary Applications," is well done. The references are extensive, and the experienced mathematical reader, by picking and choosing, can gain a good introduction to a growing body of mathematics. We feel, however, that the defects mentioned above greatly reduce the value of the work as a textbook. At a time when "relevance" is widely desired, the material presented here ought to be of particular interest, for it is indeed mathematics directly relevant to the real world. That the student may well fail to be convinced of this by the text seems to us its most serious defect.

PAUL WILLIG, Stevens Institute of Technology

Finite Probability. By Michael Gemignani. Addison-Wesley, Reading, Massachusetts, 1970. 122 pp. $1.95. (Telegraphic Review, August 1970.)

This book would seem to me to be an excellent textbook for its stated purposes, a "development of the theory of probability for finite sample spaces," and an "introduction to formal mathematical systems," requiring only basic high school algebra as a prerequisite.

The sequence of topics and manner of presentation are exemplary. The reader is prepared for a mathematical notion before it is introduced, by means of discussions either of practical considerations or of logical needs. When a new notion is introduced, it is done with care, and the student is led to see the need for rigor. "Practical" examples are carefully labelled as only possible uses or motivations rather than sole justifications.

A nice blend is given of development of a logical sequence of topics and of indications to the student that he is just scratching the surface, without overwhelming him with the enormity of what he has not yet studied.

I feel that this small volume should, in fact, give an unoriented student insight into mathematical thinking via the subject of finite probability, and should do so in a way that gains his respect and interest.

P. D. MINTON, Southern Methodist University

GPS: A Case Study in Generality and Problem Solving. By George W. Ernst. Academic Press, 1969. 307 pp. $15. (Telegraphic Review, February 1970.)

In the middle 1950s Allen Newell, Clifford Shaw, and Herbert Simon developed the "General Problem Solver" (GPS) computer program as a step toward mechanical reasoning. GPS was first given a *representation*, consisting of (a) a definition of states of the world, (b) a description of operators for moving between states, and (c) a table indicating which operators affected which differences between states. A problem was defined by stating a starting state and a goal state. The program generated a sequence of moves to go from start to goal. Since the representation is separate from the code for choosing moves, the name GPS is justified. Published descriptions of GPS's performance and internal characteristics were sketchy, but the idea had a major influence on Artificial Intelligence research.

This book, based largely on Ernst's thesis, provides the detailed description required to fill the literature gap. The description will let a programmer reconstruct a progam "in the spirit" of GPS, though he could not duplicate the code. The problems involved in the tradeoff between programming efficiency and generality are stated. There is so much detail that the forest is often lost for the trees, but since the earlier publications sketch forest only, the two together will be very helpful.

Generality is claimed for GPS by showing how it attacked twelve tasks. One of these was unsolvable, the rest are similar to the easier mathematical puzzles in *Scientific American*. This is not damaging since GPS was intended to illustrate an approach, not to be a working program. The power and deficiencies of the state space representation are well shown. I concluded that the GPS approach is a reasonable one for "uninspired" problem solving, but, as the authors point out, GPS has no mechanism for generating or evaluating representations, and thus makes no great discoveries.

Portions of the book will be read by every serious student of Artificial Intelligence. All of it will be of interest only to a few specialists. It is a valuable supplementary text for courses in Artificial Intelligence, and as such it should be in university libraries. The book would have been much improved by more careful editing, both to avoid superfluous detail and to eliminate the numerous and confusing typographical errors.

EARL HUNT, University of Washington

Probability. By Peter Whittle. Penguin, New York, 1970. 238 pp. $5.95. (Telegraphic Review, November 1970.)

This is an exceedingly interesting book to anyone with a broad background in probability. The author axiomatizes the notion of expectation rather than probability, and this makes for interesting wrinkles in the standard fabric. The appeal of this approach is three-fold (i) averages correspond more closely to what one observes than do probabilities, (ii) far less distinction need be made between sample spaces which are countable and those which are not, and (iii) one need not be familiar with any of the measure-theoretic machinery for constructing integrals out of postulated measures.

In the last chapter the author gives an elegant discussion of quantum mechanics, and a brief treatment of information theory, dynamic programming, and stochastic differential equations.

Internal referencing within the text is excellent.

Upon careful reading one finds a scope, uniqueness, and novelty infrequent in any field. This, though delightful, has its drawbacks for use as a classroom text. The prospective teacher must have more depth than is necessary to use a book such as Feller, volume I. Students cannot be prevented from reading the more difficult and necessarily incomplete sections which show the interconnection between this view and the more usual approach. Thus one skirts such major

classical corner stones as the Radon-Nikodym theorem, the construction of integrals from measures, etc.

Two additional deterrents to classroom use are a total lack of solutions to problem sets, and a weakness in references for further reading in areas where student appetite has been whetted (e.g., where may the student read further about the Ising problem or martingales?).

Even with these shortcomings one should carefully examine *Probability* by Peter Whittle before choosing a text for an introductory probability course.

R. S. PINKHAM, Stevens Institute of Technology

Qualitative Theory of Ordinary Differential Equations. An Introduction. By Fred Brauer and John A. Nohel. Benjamin, New York, 1969. xi+314 pp. $12.50. (Telegraphic Review, June-July 1970.)

This is a very good book which I would recommend to those who wish to learn the basic ideas and techniques of stability theory. It should be especially appealing to engineering, mathematics, and physics students for it motivates and exemplifies the concepts with many well-chosen examples and illustrations. In addition the book attains the level of rigor which most mathematicians demand without letting the main ideas and techniques become submerged in a mass of detail. Altogether it is a very well-written book which should satisfy all those for whom it is intended—plus some others. Which brings me to a minor criticism. The book will also be of interest to students of mathematical economics; a few well-chosen examples from economics would have helped open new horizons for some others as well as making the book more attractive to economics students.

The topics covered in the first three chapters are the standard (but useful) theorems concerning linear systems, existence and uniqueness of solutions, and dependence on parameters. The next three chapters, which are the heart of the book, deal with stability theory and its applications to oscillation phenomena, self-excited oscillations, and the regulator problem of Lurie. Since the main concept is stability it might be in order to paraphrase the definition. A solution $\phi(t)$ of $y'=f(t, y)$, $y(t_0)=y_0$, on $[t_0, \infty)$ is said to be stable iff for each $\epsilon>0$ there exists a $\delta>0$ such that if ψ is a solution of $y'=f(t, y)$ with $|\psi(t_0)-y_0|<\delta$, then ψ exists for $t\geqq t_0$ and satisfies $|\psi(t)-\phi(t)|<\epsilon$ for $t\geqq t_0$. If ϕ is stable and, in addition, every solution starting (at $t=t_0$) near y_0 approaches ϕ as $t\to\infty$, then ϕ is said to be asymptotically stable. Both the standard method of examining small perturbations from a linear system which has known stability properties, and the direct method of Lyapunov, which requires one to find an auxiliary function, are well treated. The problem of determining the size of the region of asymptotic stability is explored—a definite plus.

There are routine computations, completions of mathematical arguments, extensions of theorems, and applications to physical problems among the numerous exercises in this book. The main innovation in style comes from the

placing of the exercises not at the end of sections or chapters, but in the body of the text where they naturally occur. This placement of exercises should, if followed by the student, substantially aid in understanding the ideas and concepts contained in this book. On the other hand, this placement interrupts, to some extent, the continuity of the presentation. It will be interesting to hear the reaction of students to this innovation.

The prospective reader should know how to solve elementary differential equations, know the basic properties of continuous functions, and have a nodding acquaintance with linear algebra.

To those who wish to work and learn about stability theory, I heartily recommend this book.

F. S. VAN VLECK, University of Kansas

An Introduction to Number Theory. By Harold M. Stark. Markham, Chicago, 1970. 357 pp. \$8.50. (Telegraphic Review, June–July 1970.)

Stark's book is intended as a first course in elementary number theory. The first three chapters constitute a readable introduction to congruences, unique factorization of rational integers, the cyclicity of the units in $\mathbf{Z}/p\mathbf{Z}$ and the Chinese Remainder Theorem. One of the most attractive features of this book is the inclusion of millions (approximately) of exercises at (almost) all levels. It is in these exercises that the students find mention of better things (quadratic reciprocity, Gauss sums, etc.). Chapter 5 constitutes an introduction to diophantine equations (mostly over $\mathbf{Z}$) along traditional lines ($x^4+y^4=z^2$, Pell's equation (partial)). Congruences are used to obtain nonexistence theorems ($x^2=-1$ has no solution in $\mathbf{Z}$ because $x^2\equiv -1$ (3) has no solution) and the Pythagorean triples are classified.

The most exciting chapter, Chapter 8, is the one on quadratic number fields. Here the author's expert knowledge of the field is sparingly applied to give a lucid introduction to this subject. Elementary results are established where possible ($Q(\sqrt{d})$, $d<0$, $d\not\equiv 1(4)$ is U.F.D.$\Leftrightarrow d=-1$ or -2) and the reader is brought up to date on the situation regarding unique factorization in imaginary quadratic fields. Again the exercises are well chosen and sprinkled throughout are interesting historical comments. Chapters 1, 2, 3, 5, and 8 combine to give a good introductory course.

In addition Stark has included a chapter on magic squares, one on irrational numbers, and a pleasing account of continued fractions.

The book offers a variety of possibilities for a first course. There is at the end a useful bibliography, special references and a reminder of the Greek alphabet.

In conclusion, Stark has written a delightful leisurely account of elementary number theory with little or no ideal theoretic premeditation, included an abundant supply of great exercises, and ended with an exciting chapter on quadratic number fields.

K. F. IRELAND, Bowdoin College

TELEGRAPHIC REVIEWS

Telegraphic reviews are intended to give prompt notice of books, with sufficient information to assist in deciding whether to order an examination copy or suggest library purchase. Possible uses are indicated as follows:

B = college bookstore stock L = library purchase
P = professional reading S = supplementary reading
T = textbook E = teacher education
13 to 18 = freshman to second year graduate level usage
1 to 4 = approximate time in semesters to cover text
* = positive emphasis ? = negative emphasis

Books on high school material (pre-calculus) are denoted REMEDIAL, and normally receive telegraphic reviews only if they are written for college students. Publishers are denoted by the standard abbreviations used in *Books in Print*, which gives complete addresses.

ALGEBRA AND QUANTUM MECHANICS, P, L. *Lecture Notes in Physics, Volume 6. Group Representations in Mathematics and Physics: Battelle Seattle 1969 Rencontres.* Ed: V. Bargmann. Springer-Verlag, 1970, v + 340 pp, $6.90 (P). Valuable bridge for mathematics--theoretical physics gap. These are lectures, not polished articles, but do include presentations from both disciplines and extensive bibliographies. Provides unifying terminology and productive cross-fertilization of ideas. J.C.

ALGEBRA, LINEAR ALGEBRA, T*(13 OR 14: 1). *Linear Algebra with Applications: Including Linear Programming.* Hugh Campbell. Appleton-Century-Crofts, 1971, xiii + 441 pp, $10.95. A text designed to precede, parallel, or follow a calculus course. Applications and examples are many and varied, with applications to the calculus starred. Obviously a great deal of thought has gone into these applications with the hope that the students will find it relevant to a broad spectrum of disciplines. L.L.K.

ALGEBRA, NUMBER THEORY AND LINEAR ALGEBRA, T(14-15: 2), L(UNDERGRADUATE). *Linear Algebra.* Georgi E. Shilov. Transl. and Ed: Richard A. Silverman. P-H, 1971, xi + 387 pp, $12.95. An undergraduate text dealing almost entirely with finite-dimensional spaces, but over an arbitrary field of characteristic zero except when special results can be obtained over the real or complex fields. The pace is fairly leisurely, the spirit somewhat old fashioned (except for an appendix on categories of finite-dimensional spaces), the coverage extensive. Problems at the end of each chapter, solutions or hints for all of them at the end of the book. J.D.-B.

ALGEBRA AND NUMBER THEORY, P, L(RESEARCH). *Arithmetic Groups: Courant Institute of Mathematical Sciences.* J.E. Humphreys. NYU, 1971, iv + 121 pp, $2.75 (P). Notes on a course given at the Courant Institute in 1970 and intended as an introduction to the modern theory of arithmetic groups (A. Borel's *Introduction aux groups arithmétiques).* Only concrete examples are treated, but the methods are general. Proofs are usually only sketched, and many of the exercises scattered through the text are essential to later developments. J.D.-B.

APPLICATIONS, S*, P, L. *Applications of Mathematical Programming Techniques*. Ed: E.M.L. Beale. Am Elsevier, 1970, ix + 451 pp, $19.50. 29 interesting papers from a NATO conference held at Cambridge, England on June 24-28, 1968, varying widely from Dantzig's "Linear Programming and its Progeny" to Stolley's "Applications of Quadratic Programming to the Promotion System of Professional Officers in the German Air Force." R.W.N.

APPLICATIONS AND NETWORKS, P, L. *Lumped and Distributed Passive Networks: A Generalized and Advanced Viewpoint*. M. Ronald Wohlers. Acad Pr, 1969, 235 pp, $12. A research monograph in the Electrical Science series. The theoretical development of passive networks aided by appendices on distribution theory and the inversion of Sturm-Liouville operators and extended by one on research problems. Applications to engineering and physics. R.W.N.

APPLICATIONS AND NETWORKS, P. *Network Flow, Transportation and Scheduling: Theory and Algorithms*. Masao Iri. Acad Pr, 1969, x + 316 pp, $16.50. An inclusive coverage, from convexity and topological properties of general networks to algorithms for particular linear networks, based mainly on the author's articles in the Memoirs of the Research Association of Applied Geometry of Japan from 1956 to 1968. R.W.N.

CALCULUS, T(13: 2). *Elementary Technical Mathematics with Calculus*. Frank L. Juszli and Charles A. Rodgers. P-H, 1971, xii + 704 pp $11.95. This is mostly a pre-calculus text treating algebra, geometry, trigonometry and finally, less than 200 pages of calculus. This can only serve as a brief introduction to the calculus but may be sufficient for those who only need it as a tool. L.L.K.

CELESTIAL MECHANICS, P, L. *Celestial Mechanics. Dynamical Principles and Transformation Theory, Volume I*. Yusuke Hagihara. MIT Pr, 1970, xiii + 689 pp, $25. The first volume of a five volume treatise on celestial mechanics stimulated by the launching of artificial celestial bodies. "It is the aim of the present series of volumes to examine on rigorous mathematical grounds the question of whether the deductive process adopted in celestial mechanics can be logically justified." Part I, Dynamical Principles: principles of analytical dynamics, quasi-periodic motion, particular solutions of the many-body problems. Part II, Transformation Theory: continuous groups of transformations, reduction of the n-body problem, algebraic integrals and uniform integrals. R.B.K.

COMBINATORICS, T(16-17: 1), P, B, L. *Principles of Combinatorics. Mathematics in Science and Engineering, Volume 72*. C. Berge. Acad Pr, 1971, viii + 176 pp, $10. An introduction to the principles of combinatorial analysis. The exposition is both elegant and intuitive. A little algebra and an interest in the subject is all that is needed to follow the book. There are no problems but there are lengthy bibliographies. W.C.R.

COMBINATORICS, COMPUTER SCIENCE, T(16-17: 1), S*, P. *Elements of Combinatorical Computing*. Mark B. Wells. Pergamon Pr, 1971, xiv + 258 pp, $12. Concerned with the algorithms for efficient manipulation of combinatorial objects presented in terms of a natural programming language. Many interesting applications such as branch merging, serial numbers for numerical partitions, isomorph

rejection and a study of the four-color problem. R.W.N.

COMPLEX ANALYSIS, T(18), P, L. *The Kernel Function and Conformal Mapping, Second Edition.* Stefan Bergman. AMS, 1970, x + 257 pp, $20. The author presents a number of methods and principles which are obtained by studying the kernel function of a complete system of orthonormal complex functions. The emphasis is on the applications to the theory of conformal mappings of multiply-connected domains. Applications to the theory of functions of two complex variables are included in this revised edition. Contains a very extensive bibliography. T.A.V.

COMPUTER SCIENCE, P, S, L. *Lecture Notes in Mathematics-188: Symposium on Semantics of Algorithmic Languages.* Ed: E. Engeler. Springer-Verlag, 1971, vi + 372 pp, $7.60 (P). The results of a write-in symposium: 14 current papers on topics such as logical and algebraic foundations, proving correctness and program synthesis. R.W.N.

EDUCATION, ELEMENTARY, E(ELEMENTARY). *Geometry for Elementary Teachers.* John E. Young and Grace A. Bush. Holden-Day, 1971, xii + 273 pp, $9.95. Makes an attempt to incorporate intuitive ideas into a rerun of Euclidean geometry. But, in general, it fails to demonstrate how this formally developed geometry can be presented to elementary students. J.N.C.

*GEOMETRY, E, S, L. *An Introduction to Transformational Geometry.* Frank M. Eccles. A-W, 1971, v + 177 pp, $2.40 (P). This book is primarily intended to present transformational geometry to the high school student following a compressed treatment of a standard Euclidean geometry course and "to bring the student back into the Mathematical Mainstream by placing the geometry in an algebraic context." Well-written, with interesting applications and a variety of good problems. It would also be an excellent supplementary text for the future secondary teacher. J.N.C.

GEOMETRY, S, P, L. *Twisted Honeycombs. Conference Board of the Mathematical Sciences Regional Conference Series in Mathematics, Number 4.* H.S.M. Coxeter. AMS, 1970, iv + 47 pp, $2.70 (P). A series of lectures sponsored by NSF and the University of Maine. J.N.C.

HARMONIC ANALYSIS, P, L. *Fourier Analysis on Matrix Space. Memoirs of the American Mathematical Society, Number 108.* Stephen S. Gelbart. AMS, 1971, 77 pp, $1.90 (P). Treats two related problems of harmonic analysis on n x n real matrix space, the first the decomposition of the additive Fourier operator with respect to the group representation theory of G = GL(n, R), the second analytical continuation and functional equation of certain zeta-functions defined on G. T.A.V.

HISTORY, T(13-16), S, L. *History of Binary and Other Nondecimal Numeration.* Anton Glaser. Anton Glaser (1237 Whitney Road, Southampton, Pa, 18966). 1971, ix + 196 pp, $4 (P). Based on the author's doctoral dissertation, this book contains a lot of material about a fairly narrow topic. Ranges from pre-Leibnitz to applications in modern computers. Bibliography good. A.G.

NUMERICAL ANALYSIS, P, L. *Iterationsverfahren Numerische Mathematik Approximationstheorie.* International Series of Numerical Mathematics. Birkhauser Verlag Basel, 1970, 257 pp, 8.57. The 24 papers presented at the following conferences: Nonlinear Problems in Numerical Mathematics (Nov. 17-23, 1968), Numerical Methods in Approximation Theory (June 8-14, 1969), and Iteration Methods in Numerical Mathematics (Nov. 16-22, 1969); all at the Mathematischen Forschungs Institut of Oberwolfach. A few of these are survey papers. Most are specialized with excellent bibliographies. R.W.N.

NUMERICAL ANALYSIS AND APPLICATIONS, OPERATIONS RESEARCH, L, P. *Integer Programming. Mathematics in Science and Engineering, Volume 76.* Harold Greenberg. Acad Pr, 1971, xii + 196 pp, $11.50. Linear programming techniques in problems restricted entirely to integers. Summary of applications, review of linear programming. Solutions by all-integer methods, efficient enumeration, modifying continuous solutions, dynamic programming, and branch and bound procedures. Many references. J.G.L.

NUMERICAL ANALYSIS, PARTIAL DIFFERENTIAL EQUATIONS, S, P*, L. *Numerical Solution of Partial Differential Equations-II. Synspade 1970.* Ed: Bert Hubbard. Acad Pr, 1971, ix + 649 pp, $14. This volume contains the 18 invited papers delivered at the second symposium on the numerical solution of partial differential equations which was held May 11-15, 1970 at the University of Maryland. The papers vary from the mathematical analyses to the applications of numerical methods for partial differential equations. Several are on the finite element method. R.W.N.

NUMERICAL METHODS, S, P, L(RESEARCH). *Automatic Programming, Numerical Methods and Functional Analysis. Proceedings of the Steklov Institute of Mathematics. Number 96.* Ed: V.N. Faddeeva. AMS, 1970, vii + 323 pp, $22.40 (P). Translated by Leo Ebner. The 6 papers dealing with the development of Kantorovich's automatic programming system reveal many differences from similar developments elsewhere. The 9 papers on numerical analysis and 2 on functional analysis are much more diverse. R.W.N.

OPERATOR THEORY(NONLINEAR), P, L. *Nonlinear Functional Analysis and Applications: Proceedings of an Advanced Seminar Conducted by the Mathematics Research Center, The University of Wisconsin, Madison, October 12-14, 1970.* Ed: Louis B. Rall. Acad Pr, 1971, vii + 586 pp, $11. A useful collection of papers of interest to the novice and expert alike. The first two are introductory. L. Collatz, "Some Applications of Functional Analysis to Analysis, Particularly to Nonlinear Integral Equations"; R.A. Tapia, "The Differentiation and Integration of Nonlinear Operators"; M.Z. Nashed, "Differentiability and Related Properties of Nonlinear Operators: Some Aspects of the Role of Differentials in Nonlinear Functional Analysis"; M.Z. Nashed, "Generalized Inverses, Normal Solvability, and Interation for Singular Operator Equations"; Patricia M. Prenter, "On Polynomial Operators and Equations"; James W. Daniel, "Applications and Methods for the Minimization of Functionals"; J.E. Dennis, Jr., "Toward a Unified Convergence Theory for Newton-Like Methods"; David L. Russell, "Operator Solutions of Nonlinear Equations in Optimal Control Problems"; Peter D. Robinson, "Complementary Variational Principles". Nashed's paper on differentials has 363 usefully

organized references. R.B.K.

PARTIAL DIFFERENTIAL EQUATIONS, FOURIER ANALYSIS, T(1), L. *Introduction to Partial Differential Equations from Fourier Series to Boundary-value Problems.* Arne Broman. A-W, 1968, x + 179 pp, $12.95. An attractive introduction to Fourier series, orthogonal systems, orthogonal polynomials, Fourier transforms, Laplace Transforms, Bessel Functions, partial differential equations of the first order, and partial differential equations of the second order. Designed to follow introductory courses in ordinary differential equations and complex variables, but not an awful lot depends upon the latter. R.B.K.

PRE-CALCULUS/ALGEBRA, T***(13: 2). *Algebraic Elementary Functions and Relations.* Donald R. Horner. HR & W, 1971, xii + 428 pp, $9. A well-written text for students needing a careful treatment of precalculus topics: sets, functions, induction, relations, real number system, algebraic functions, applications. Many exercises and solutions. A.G.

PROBABILITY AND STATISTICS, P. *Selected Translations in Mathematical Statistics and Probability, Volume 9.* AMS, 1971, vi + 315 pp, $16.20. Translations of 29 papers (28 from the Russian) which first appeared between 1957 and 1967. Most of them are concerned with stochastic processes. F.L.W.

PROBABILITY AND STATISTICS, T(18: 2), P. *Brownian Motion and Diffusion.* David Freedman. Holden-Day, 1971, xii + 231 pp, $12.95. This text, with the authors' *Markov Chains* and *Approximating Countable Markov Chains*, completes a trilogy on Markov processes. A constructive approach is emphasized. The necessary probability theory and analysis are reviewed in an appendix. F.L.W.

PROBABILITY AND STATISTICS, S, P, L. *Statistical Tolerance Regions, Classical and Bayesian. Griffin's Statistical Monographs and Courses, Number 26.* Irwin Guttman. Hafner, 1970, ix + 150 pp, $7.95 (P). General discussion and survey of the literature on tolerance regions from the classical (80 pages) and the Bayesian (20 pages) points of view. 45 pages of tables. F.L.W.

PROBABILITY AND STATISTICS, T(18: 2), S, P, L. *Nonparametric Methods in Multivariate Analysis.* Madan Lal Puri and Pranab Kumar Sen. Wiley, 1971, xi + 440 pp, $19.95. Presupposes (but reviews) abstract probability theory and univariate non-parametric statistics. Includes multivariate tests of location and scale, point and interval estimation in linear models, multifactor designs, testing independence of different subsets of a stochastic vector, and testing the identity of dispersion matrices. Many exercises. Extensive bibliography. F.L.W.

PROBABILITY AND STATISTICS, S, P, L*. *Statistical Methods for Research Workers.* 14th Edition. Sir Ronald A. Fisher. Hafner, 1970, xiii + 362 pp, $4.95 (P). Presumably the last edition of one of Fisher's classics. Published posthumously, it contains only minor changes from the previous edition. R.S.K.

PROBABILITY AND STATISTICS, P, S. *Advances in Probability and*

Related Topics, Volume I. Ed: Peter Ney. Marcel Dekker, 1971, xii + 215 pp, $12.50. Papers on random walks and discrete subgroups of Lie groups (by Harry Furstenberg), using random measures and random sets in harmonic analysis (by J.P. Kahane), finite Toeplitz operators (by I.I. Hirschmann, Jr.), and an introduction to matching theory (by Gian-Carlo Rota and L.H. Harper). F.L.W.

PROBABILITY AND STATISTICS, T(16-17: 1, 2), S, P, L. *Advanced Statistical Methods in Biometric Research*. G. Radhakrishna Rao. Hafner, 1970, xvii + 390 pp, $15. A nice blend of theory and numerical examples. Includes chapters on vectors and matrices, distribution theory, linear estimation and tests of hypotheses, large sample methods, tests of homogeneity, multivariate analysis, classification problems, and group constellations. No exercises except for a very few in an appendix. F.L.W.

PROBABILITY AND STATISTICS, S, P, L. *Experimental Design: Selected Papers of Frank Yates*. Frank Yates. Hafner, 1970, xi + 296 pp, $13.95. "A representative selection of Yates's publications on experimental design with particular attention to some of the earlier papers now not readily accessible to statisticians." Author's notes put the papers in historical perspective. A complete bibliography of his works is included. F.L.W.

PROBABILITY, ERGODIC THEORY, P. *Lecture Notes in Mathematics-160: Contributions to Ergodic Theory and Probability*. Ed: A.Dold and B. Eckmann. Springer-Verlag, 1970, vii + 277 pp, $6.80 (P). Proceedings of the First Midwestern Conference on Ergodic Theory held at the Ohio State University, March 27-30, 1970. Papers by Anatole Beck, A. Brunel, R.V. Chacon, R.V. Chacon and T. Schwartzbauer, P. Erdös and A. Rényi, Arshag B. Hajian and Shizuo Kakutani, D.L. Hanson, S. Horowitz, A. Ionescu Tulcea, C. Ionescu Tulcea, Konrad Jacobs, Eugene M. Klimko, Ulrich Krengel, Wolfgang Krieger, Donald S. Ornstein, Donald S. Ornstein and Louis Sucheston, Fredos Papangelou, M.M. Rao, and Thomas R. Terrell. R.B.K.

REAL ANALYSIS, T*(16-17: 1), B, L. *Measure and Integration, Second Edition*. M.E. Munroe. A-W, 1971, xii + 290 pp, $12.50. Differs from the popular first edition by the addition of a chapter on functional analysis including the Hahn-Banach, Banach-Steinhaus and closed graph theorems along with a discussion of weak and weak* convergence. T.A.V.

REAL VARIABLES, INEQUALITIES, P, L*. *Analytic Inequalities*. D.S. Mitrinovic. *Die Grundlehren der mathematischen Wissenschaften, Band 165*. Springer-Verlag, 1970, xi + 400 pp, $27.30. A fascinating book on inequalities in analysis. After an introduction to convex functions of one variable, general inequalities on n-space are presented, including generalizations of the well-known inequalities. The last half of the book is a collection of over 450 special inequalities, many from problems in the Monthly. All results are referenced. Over 750 names are cited, some several times. R.B.K.

REMEDIAL, T(13: 1). *Modern Analytic Geometry*. Gerald C. Preston and Anthony R. Lovaglia. Har-Row, 1971, x + 319 pp, $7.95. This book may be attractive to those who prefer a thorough treatment of

analytic geometry and an introduction to linear algebra before starting calculus. Exercises are numerous and diversified. K.W.

*REMEDIAL, S, L. *Geometry: A Transformation Approach.* Arthur F. Coxford and Zalman P. Usiskin. Laidlaw Brothers, 1971, xii + 612 pp, $5.22. An attractive, well-written high school text which presents the standard content of Euclidean geometry via transformations--its publication should be noted by the college community. J.N.C.

REMEDIAL. *Mathematics Through Statistics.* Louis Auslander, *et. al.* Aldine-Atherton, Inc., 1971, 388 pp, $4.95 (P). This book is designed for the student who finds mathematics so difficult that he hasn't progressed beyond sixth grade material. There are 388 pages and on page 300 this problem can be found: "At basketball practice Jon made 3 baskets out of 7 tries. What percent of baskets did he make?" Normal distribution and probability are handled on a very elementary level. L.L.K.

REMEDIAL, T(1), S, AND EDUCATION(ELEMENTARY). *Arithmetic: A Modern Approach.* Mervin L. Keedy and Marvin L. Bittinger. A-W, 1971, xii + 420 pp, $5.75 (P). As the authors say "this book contains... features not usually found in a college textbook" -- basic arithmetic. No index. J.N.C.

REMEDIAL AND PRE-CALCULUS, T(13: 1, 2), S, L. *Algebra and Trigonometry.* William J. Bruce and Edgar Phibbs. Appleton-Century-Crofts, 1971, xv + 446 pp, $9.95. Has greater coverage than the usual text at this level. Some matrix algebra comes in at an early stage. There is an extensive chapter on operations in linear algebra. K.W.

REMEDIAL, BUSINESS MATHEMATICS, T(13). *Elementary Business Mathematics.* George F. Hadley. Richard D. Irwin, 1971, xii + 682 pp, $10.95. Straightforward treatment on a remedial/elementary level of: arithmetic, discounts, markups, payroll, taxes, interest, insurance, graphs, statistical variation, and linear programming.A.G.

REMEDIAL, ENGINEERING, ?. *An Approach to Engineering Mathematics.* Arthur H. Douglas. Pergamon Pr, 1971, viii + 143 pp, $4 (P). "An attempt to re-state in simple terms, and in logical sequence, some basic principles and techniques which the ordinary engineer uses in his work, its object is to guide rather than instruct." The language is so simple that either guidance or instruction is doubtful. Example: "The term function is used here to denote any useful expression built up from arithmetical or algebraic numbers." R.B.K.

REMEDIAL, GRADE SCHOOL. *Beginning Mathematics for College Students.* John D. Leonard and Blaine A. Warner. P-H, 1971, 389 pp, $6.95 (P). An arithmetic workbook, place value to percentage (with pages of drill on 2 + 2 = ? and other Basic Number Facts). Adults who want to learn arithmetic have a legitimate need for instruction and instructional materials (and this seems an attractively done if standard treatment). It is neither necessary nor possible for this book to be used for a course "at the college level." L.A.S.

REMEDIAL, JR. HIGH. *A Programmed Course in Basic Algebra.* William C. Beck and James R. Trier. A-W, 1971, 458 pp, $5.95 (P). For

students with a year or less of high school algebra. L.A.S.

REMEDIAL, TRIGONOMETRY, T(13: 1). *Essentials of Trigonometry.* Irving Drooyan, Walter Hadel, and Charles C. Carico. Macmillan, 1971, ix + 336 pp, $8.95. Remedial. Usual stuff on solving triangles, trigonometry functions, identities, and complex numbers. Appendices cover logarithms, special formulas for solving triangles, tables of trigonometry and logarithmic functions, and axioms for the real numbers (missing the completeness axiom). Attractive format. Old trig in new clothes. A.G.

SEVERAL COMPLEX VARIABLES, P, L. *Lecture Notes in Mathematics-184: Symposium on Several Complex Variables, Park City, Utah, 1970.* Ed: R.M. Brooks. Springer-Verlag, 1971, 234 pp, $5.80 (P). Papers based on talks given at the Symposium held March 30-April 3, 1970, by F. Birtel and W. Zame, D.D. Clayton, Michael Freeman, T.W. Gamelin, John Wermer, D.C. Spencer, Wilhelm Stoll, R.O. Wells, Jr., Yum-Tong Siu, Andrew Markoe, Andrew Markoe and Hugo Rossi, Bernard Shiffman, and Robert O. Kujala. R.B.K.

SEVERAL COMPLEX VARIABLES, P. *Lecture Notes in Mathematics-155: Several Complex Variables I, Maryland 1970.* Ed: A. Dold and B. Eckmann. Springer-Verlag, 1970, 214 pp, $5.30 (P). *Lecture Notes in Mathematics-185: Several Complex Variables II, Maryland 1970.* Ed: John Horváth. 1971, 287 pp, $6.60 (P). Proceedings of the International Mathematical Conference, held at College Park, April 6-17, 1970. Volume II papers by Phillip Griffiths, R.C. Gunning, Michel Hervé, Jun-ichi Igusa, Tomio Kubota, Pierre Lelong, H.L. Royden, Ichiro Satake, Goro Shimura, Yum-Tong Siu and Joseph A. Wolf. Volume I papers by Walter L. Bailey, Lipman Bers, Enrico Bombieri, M. Eichler, Hans Grauert and Oswald Riemenschneider, M. Herrera, Adam Korányi, H. Maass, Yozo Matsushima, Raghavan Narasimhan, I.I. Pjateckii-Šapiro, Wilhelm Stoll, Walter Thimm, and Edkardo Vesentini. R.B.K.

STATISTICAL PHYSICS, P. *Statistical Mechanics at the Turn of the Decade.* E.G.D. Cohen. Marcel Dekker, 1971, viii + 235 pp, $12.50. Record of a symposium commemorating the seventieth birthday of Professor G.E. Uhlenbeck held at Northwestern in October, 1969. Contributions by E.G.D. Cohen, C. Domb, Freeman J. Dyson, P.C. Hohenberg, Paul C. Martin, D. Ruell, A.J.F. Siegert, and A.S. Wightman. R.B.K.

TOPOLOGY, T(16-17: 1 OR 2), L. *Cellular Decompositions of 3-Manifolds that Yield 3-Manifolds. Memoirs of the American Mathematical Society, Number 107.* Steven Armentrout. AMS, 1971, 72 pp, $1.90 (P). Tersely written memoir. A.G.

Reviewers Whose Initials Appear Above

Clarence Carlson, St. Olaf; James Cederberg, St. Olaf; Judith N. Cederberg, St. Olaf; John Dyer-Bennet, Carleton; Arthur Gropen, Carleton; Lorraine L. Keller, St. Olaf; Roger B. Kirchner, Carleton; Richard S. Kleber, St. Olaf; John G. Lewis, St. Olaf; R.W. Nau, Carleton; William C. Ramaley, Carleton; Linda A. Seebach, St. Olaf; T.A. Vessey, St. Olaf; Kenneth Wegner, Carleton; Frank L. Wolf, Carleton.

NEWS AND NOTICES

EDITED BY RAOUL HAILPERN, SUNY at Buffalo

Readers are invited to contribute to the general interest of this Department by sending news items to Mathematical Association of America, 1225 Connecticut Avenue, N.W., Washington, D. C. 20036. Items must be submitted at least two months before publication can take place.

PERSONAL ITEMS

Professor C. D. Harris, University of Tulsa, represented the Association at the inauguration of R. E. Collier as President of Northeastern State College, Tahlequah, on May 6, 1971.

Professor C. H. Heinke, Capital University, was honored with a 25 Year Service Award by the Capital University Alumni Association.

Professor Mina S. Rees, President of the Graduate Division of the City University of New York and 1971 President of the American Association for the Advancement of Science, was awarded an honorary Doctor of Science degree from the University of Rochester.

Dr. J. B. Rosser, Professor of Mathematics and Computer Sciences and Director of the Mathematics Research Center on the Madison Campus of the University of Wisconsin, received an honorary Doctor of Science degree from Otterbein College.

Professor M. E. Wick, Wisconsin State University, Eau Claire, represented the Association at the inauguration of K. E. Lindner as President of Wisconsin State University, La Crosse, on April 23, 1971.

Professor Martin Wright, University of Houston, represented the Association at the inauguration of E. T. Powers as President of Sam Houston State University on April 27, 1971.

Assistant Professor Syed Asadula, Saint Francis Xavier University, Antigonish' Nova Scotia, has been promoted to Associate Professor.

Dr. R. A. Melter, University of South Carolina, has been appointed Associate Professor at Southampton College of Long Island University.

Dr. Nagendra Pandey, Oregon State University, has been appointed Assistant Professor at Montana College of Mineral Science and Technology, Butte.

Professor W. A. Albrecht, Jr., California State College at Long Beach, died on April 16, 1971 at the age of 51. He was a member of the Association for thirteen years.

Professor C. H. Butler, Western Michigan University, died on November 21, 1970 at the age of 76. He was a member of the Association for twenty-nine years.

Professor Emeritus Walter B. Ford, University of Michigan, died on February 24, 1971 at the age of 96. A Charter Member of the Association, Professor Ford was Editor of the MONTHLY from 1922 to 1927 and President of the Association from 1927 to 1928.

Dr. A. G. Swanson, Gustavus Adolphus College, died on January 30, 1971 at the age of 64. He was a member of the Association for forty-two years.

MATHEMATICAL ASSOCIATION OF AMERICA

Official Reports and Communications

APRIL MEETING OF THE METROPOLITAN NEW YORK SECTION

The thirtieth annual meeting of the Metropolitan New York Section of the MAA was held on April 3, 1971, at Nassau Community College. A total of 86 people were in

attendance. Dr. A. E. Donor, Associate Dean of Instruction at Nassau Community College, welcomed the group.

Professor Mary Dolciani of Hunter College, Chairman of the Section, presided at the morning session. The following talks were given:

1. *Associativity*, by Alex Heller, Graduate Division of the City University of New York.
2. *What is uniformization?*, by Lipman Bers, Columbia University.

Professor M. J. Hellman of Long Island University, Vice-Chairman for Senior Colleges, presided at the afternoon session which began with a panel dicussion of "Impact of Open Admissions: Preliminary Report," moderated by Professor Hellman. The panelists were Professor Edward Carroll, New York University; Professor Vicki Chuckrow, City College of New York; Mr. Gerald Lieblich, Bronx Community College.

The report of the Sectional Governor, Professor Gerald Freilich of City College of New York, is to be mailed to all present at the meeting. Professor Erwin Just of Bronx Community College and Professor Maurice Nadler of Pace College served on a panel presenting the pros and cons, respectively, of accreditation and/or certification in mathematics. Ballots with questions dealing with the adoption of such a policy were distributed among the membership present.

The activities of the Speakers' Bureau were summarized by Professor James Eastham of Queensborough Community College. Professor Howard Kleiman of Queensborough Community College presented the report of the Math. Fair Committee in the absence of its chairman, Mr. Harry Ruderman of Hunter High School. The third annual Math. Fair took place on March 21 and 28, 1971 at Pace College. Seventy students from sixteen high schools in New York City and Westchester County competed. Mr. Samson Gruber, a student at Bronx High School of Science and a gold medal winner at the Math. Fair, summarized his project which was entitled "Some Properties of Uniformly Continuous Maps." Mr. Aaron Shapiro of Midwood High School gave the Treasurer's Report. In the absence of Mr. Morton Friedman of Brooklyn College, Chairman of the Contest Committee, Mr. Shapiro summarized the 1971 results of the Metropolitan New York Section MAA Contest. The following emerged as the top high schools: Freeport, Bronx Science, Stuyvesant, Brooklyn Technical, and Martin Van Buren.

The following officers were elected for the term 1971–1973: Chairman, M. J. Hellman, Long Island University; Vice-Chairman for Colleges, Israel Rose, Lehman College; Vice-Chairman for Two-Year Colleges, Howard Kleiman, Queensborough Community College; Vice-Chairman for High Schools, Ira Ewen, John Dewey High School; Secretary, Rora Iacobacci, St. John's University; Treasurer, Aaron Shapiro, Midwood High School.

GERMANA GLIER, *Secretary*

APRIL-MAY MEETING OF THE OHIO SECTION

The fifty-fifth annual meeting of the Ohio Section was held at Ohio Wesleyan University, Delaware, Ohio, on Friday, April 30, and Saturday, May 1, 1971. Professor B. J. Yozwiak, Chairman of the Section, presided at the business meeting, and Professors Richard Laatsch, Ray Rolwing, and Stanley Dice of the Program Committee presided at the program sessions. One hundred forty-nine persons registered in attendance, including one hundred thirty members of the association.

The following officers were elected: Chairman, Professor Elwood Bohn, Miami University; Chairman Elect, Professor Robert Roberts, Denison University; Secretary-Treasurer, Professor Foster Brooks, Kent State University; Program Committee: Chairman, Professor R. H. Rolwing, University of Cincinnati, Professor S. C. Dice, Wittenberg University, Professor J. A. Murtha, Marietta College.

The following program was presented:

1. *Applications of undergraduate mathematics*, by Ben Noble, U. S. Army Mathematics Research Center and Oberlin College (invited address).
2. *Some applications of algebra to music*, by Donald Koehler, Miami University.
3. *Tire contour integrals*, by Gus Mavrigan, Youngstown State University.
4. *Error-correcting codes*, by G. G. Gilbert, Miami University.
5. *A simple method for deriving formulas for the sum of k-th powers of integers*, by Josef Blass, Bowling Green State University.
6. *An application of decision theory*, by Becky Klemm (student), Miami University.
7. *Cubical polyhedra and homotopy*, by Josef Blass, Bowling Green State University, and W. Holsztynski, Ann Arbor.
8. *A direct-table test for associativity*, by L. D. Rodabaugh, University of Akron.
9. *Ovals in finite hyperbolic planes*, by S. C. Saxena, University of Akron.
10. *An example and a question in set theory*, by D. J. Johnson, Air Force Institute of Technology.
11. *Belgian mathematics and pedagogy of Papy*, by J. L. Smith, Muskingum College.
12. *What you always wanted to know about summability, but were not interested enough to ask*, by B. J. Yozwiak, Youngstown State University (Chairman's Address).
13. *An analytic introduction to the circular functions*, by D. G. Shumway, Bowling Green State University.
14. *An examination of the trigonometric functions defined on a general ellipse and on a general hyperbola*, by Robert French (student), Miami University.
15. *Flexible scheduling for the calculus*, by H. M. MacNeille, Case Western Reserve University.
16. *Convex sets in geometry and analysis*, by Victor Klee, University of Washington, President of the Association (invited address).

FOSTER BROOKS, *Secretary*

APRIL-MAY MEETING OF THE WISCONSIN SECTION

The annual meeting of the Wisconsin Section of the MAA was held at Ripon College on April 30 and May 1, 1971. Chairman W. B. White, University of Wisconsin, Sheboygan campus, presided. Approximately 150 persons attended. After registration on Friday afternoon, April 30, the following papers were presented:

1. *Scheduling regular airlines*, by G. Uebe, University of Wisconsin, Madison, Mathematics Research Center.
2. *On univalent functions in the unit disk*, by G. M. Shah, University of Wisconsin, Waukesha.
3. *Compact spaces and products of finite spaces*, by J. D. Harris, Marquette University, Milwaukee.
4. *Mathematical films—current and prospective activity*, by R. G. Long, Lawrence University, Appleton.

After dinner on Friday evening several films were shown including "The Peano Curve", "Unsolved Problems in Geometry (Part I)", "Caroms", "Limit", and "Infinite Acres."

The Saturday morning session was opened by President Bernard Adams of Ripon College. The following papers were then presented:

1. *Figure Construction in Lumbee Geometry*, by R. Najor, Wisconsin State University, Whitewater.
2. *The non-math student encounters mathematics*, by A. Christenson, Milwaukee Area Technical College, Milwaukee.
3. *Constructions of continuous functions between continua in E^n*, by J. Sobota, Wisconsin State University, La Crosse.
4. *The computer-oriented calculus course*, by M. Engert, Wisconsin State University, Whitewater.

5. *Paracompactifications*, by J. D. Wine, Wisconsin State University, La Crosse.
6. *Repetitive play in games*, by J. Van Ryzin, University of Wisconsin, Madison.

During the business meeting which followed, Dr. John Teska was elected Chairman, Dr. Roland Christensen was elected Vice-Chairman, and Dr. Ray Wagner was elected Secretary-Treasurer.

After a luncheon break, Philip Bender of Marquette University discusssed a survey of the Section on "Accreditation and Certification". The Saturday afternoon session concluded with the presentation of the following papers:

1. *Linear programming—an introductory example*, by A. Berry, Lawrence University, Appleton.
2. *Graphs and matroids: an expository talk*, by R. Brualdi, University of Wisconsin, Madison.

R. W. CHRISTENSEN, *Secretary-Treasurer*

MAY MEETING OF THE ALLEGHENY MOUNTAIN SECTION

The Spring Meeting of the Allegheny Mountain Section of the MAA was held at Geneva College in Beaver Falls, Pennsylvania, on May 7 and 8, 1971. The meeting was held in conjunction with the Pennsylvania Council of Teachers of Mathematics.

The following two sessions of short papers were held on Friday evening:

Session on Computer Science; presider: Orrin Taulbee, University of Pittsburgh.

1. *Information theory of computer science students*, by William Conner, University of Pittsburgh.
2. *The use of computers in mathematics education to stimulate creativity*, by Thomas Dwyer, University of Pittsburgh.
3. *A student-directed computer program to explore elementary set theory*, by Frank Wimberly, University of Pittsburgh.
4. *Functions and machines*, by Preston Hammer, Pennsylvania State University.
5. *Multilevel systems, category theory, and structuring computer science courses*, by David Rine, West Virginia University.

Session in Mathematics; presider: William Beck, Chatham College.

1. *H-classes in a semigroup*, by J. B. Kim, West Virginia University.
2. *Product preserved properties of partially ordered sets*, by Charles Getchell, Lycoming College.
3. *Leibniz's extravagant hopes for binary arithmetic as a useful research tool*, by Dr. Anton Glaser, Pennsylvania State University.
4. *An attempt at characterizing open PL manifolds*, by A. J. Machusko, California State College.
5. *Some elementary applications of Dirac distribution*, by Jagdish Agrawal, California State College.

Invited addresses were presented as follows: "Key Ideas in a Modern Mathematics Program," by Dr. Harry Ruderman, Hunter College, New York; "Recent Developments in the Theory of Partition Identities," by Dr. Henry Alder, Secretary, MAA; "The Newer Mathematics for the 1970's," by Dr. Cletus Oakley, Villanova College.

A panel, "Accreditation and Certification in Mathematics," included Daniel Finkbeiner of Kenyon College as moderator with James Bartoo of Pennsylvania State University, J. C. Eaves of West Virginia University, presenting various aspects of the topic, and Charles Fuget of Indiana University of Pennsylvania presenting the accreditation program as instituted by the American Chemical Society.

Professor A. B. Cunningham, Chairman of the Allegheny Mountain Section, presided at the business meeting. Professor Frank Kocher of Pennsylvania State and Betty Miller of West Virginia reported on the High School Mathematics Contests for

Western Pennsylvania and West Virginia, respectively. Putnam Examination winners for the Section were recognized. A committee is to be appointed to develop special programs for the Section. Officers elected for a two-year term included J. C. Eaves of West Virginia University as Chairman, and Charles Cable of Allegheny College as Second Vice-Chairman. William Beck of Chatham College is the first Vice-Chairman and Melvin Woodard of Indiana University of Pennsylvania is the Secretary-Treasurer.

MELVIN WOODARD, *Secretary-Treasurer*

MAY MEETING OF THE NEW JERSEY SECTION

A joint meeting of the New Jersey Section of the MAA and the Association of Mathematics Teachers of New Jersey was held at Jersey City State College on May 1, 1971. Dr. S. L. Greitzer, Chairman of the Section, presided at the morning session, and Mr. A. L. Collard presided at the afternoon session. Sixty-three persons attended the meeting including forty members of the MAA.

During the morning session the following papers were presented:

1. *Mathematics and gerontology*, by Dr. Albert Socol, Ocean County College.
2. *The challenge of teaching mathematics in a community college*, by Professor Helen Bourgeois, County College of Morris.

The afternoon session was concluded with: "Are we Teaching Calculus in the High School?", by Mr. Ronald Zink, Mathematics Supervisor, Bricktown, New Jersey. After a brief presentation the speaker presided over a lively group discussion on various aspects of the topic.

J. K. RECKZEH, *Secretary*

MAY MEETING OF THE NORTH CENTRAL SECTION

The annual spring meeting of the North Central Section of the MAA was hosted by the School of Mathematics, University of Minnesota, Minneapolis, Minnesota, on May 8, 1971. Professor Edgar Reich, University of Minnesota, presided at the morning session and Professor Alfred Aeppli, University of Minnesota, presided at the afternoon session. One hundred fifteen persons attended the meeting, ninety of whom were members.

The following officers were elected for the 1971–72 academic year: Chairman, Kenneth Wegner, Carleton College; Chairman-Elect: W. S. Loud, University of Minnesota; Secretary-Treasurer: H. M. Anderson, Gustavus Adolphus College; Executive Committee members: W. A. Dolid, Metropolitan State Junior College, and Ernest Stennes, St. Cloud State College.

P. J. Hilton, University of Washington (Seattle) and Battelle Seattle Research Center, gave the invited address: "The Language of Categories and Functors."

Other papers presented were:

1. *An orientation and training program for teaching assistants*, by S. K. Grosser, University of Minnesota.
2. *Matrices of Fibonacci numbers—a new way to some old results*, by G. E. Bergum, South Dakota State University.
3. *Minnesota performance in the MAA High School Competition*, by Wayne Roberts, Macalester College.
4. *Utilization of the continuous right inverse in solving ordinary differential equations*, by D. K. Cohoon, University of Minnesota.
5. *Marginal subgroups*, by T. K. Teague, Gustavus Adolphus College.
6. *An alternate approach to the undergraduate curriculum*, by H. B. Coonce, Mankato State College.

WARREN THOMSEN, *Secretary-Treasurer*

CALENDAR OF FUTURE MEETINGS

Fifty-fifth Annual Meeting, Las Vegas, Nevada, January 19–21, 1972.

Fifty-third Summer Meeting, Dartmouth College, Hanover, August 28–30, 1972.

The following is a list of the Sections of the Association with dates of future meetings so far as they have been reported to the Editorial Director.

ALLEGHENY MOUNTAIN, Pennsylvania State University, Altoona, May 5–6, 1972.

FLORIDA, Central Florida Junior College, Ocala, March 17–18, 1972.

ILLINOIS, Lake Forest College, Lake Forest, May 12–13, 1972.

INDIANA, Earlham College, Richmond, October 30, 1971.

IOWA, University of Iowa, Iowa City, April 28, 1972.

KANSAS, Washburn University, Topeka, March 24–25, 1972.

KENTUCKY, Georgetown University, Spring 1972.

LOUISIANA-MISSISSIPPI, Millsaps College, Jackson, Mississippi, February 18–19, 1972.

MARYLAND-DISTRICT OF COLUMBIA-VIRGINIA

METROPOLITAN NEW YORK

MICHIGAN, Oakland University, Rochester, May 5–6, 1972.

MISSOURI, Stephens College, Columbia, May 5–6, 1972.

NEBRASKA, University of Nebraska at Omaha, Omaha, April 21–22, 1972.

NEW JERSEY, Stevens Institute of Technology, Hoboken, November 13, 1971.

NORTH CENTRAL

NORTHEASTERN, Wellesley College, Wellesley, Massachusetts, November 27, 1971.

NORTHERN CALIFORNIA, California State College at Hayward, February 5, 1972.

OHIO, Ashland College, Ashland, November 5–6, 1971.

OKLAHOMA-ARKANSAS, State College of Arkansas, Conway, Arkansas, March 10–11, 1972.

PACIFIC NORTHWEST, University of Washington, Seattle, June 16–17, 1972.

PHILADELPHIA, Lafayette College, Easton, November 20, 1971.

ROCKY MOUNTAIN, The Colorado School of Mines, Golden, May 5–6, 1972.

SOUTHEASTERN, Samford University, Birmingham, Alabama, March 24–25, 1972.

SOUTHERN CALIFORNIA, California Institute of Technology, Pasadena, March 11, 1972.

SOUTHWESTERN, University of New Mexico, Albuquerque, Spring 1972.

TEXAS, Southwest Texas State University, San Marcos, April 1972.

UPPER NEW YORK STATE, State University College at Geneseo, New York, November 6, 1971.

WISCONSIN, Wisconsin State University, Stevens Point, April 28–29, 1972.

FUTURE MEETINGS OF OTHER ORGANIZATIONS

AMERICAN ASSOCIATION FOR THE ADVANCEMENT OF SCIENCE, Philadelphia, December 26–31, 1971.

AMERICAN MATHEMATICAL SOCIETY, Las Vegas, Nevada, January 17–20, 1972.

AMERICAN SOCIETY FOR ENGINEERING EDUCATION

ASSOCIATION FOR COMPUTING MACHINERY, Boston, Massachusetts, August 14–16, 1972.

ASSOCIATION FOR SYMBOLIC LOGIC

CENTRAL ASSOCIATION OF SCIENCE AND MATHEMATICS TEACHERS, Detroit, Michigan, November 18–20, 1971.

FIBONACCI ASSOCIATION, College of the Holy Names, Oakland, California, November 13, 1971.

INSTITUTE OF MATHEMATICAL STATISTICS

MU ALPHA THETA

NATIONAL COUNCIL OF TEACHERS OF MATHEMATICS, Chicago, Illinois, April 16–19, 1972.

OPERATIONS RESEARCH SOCIETY OF AMERICA, Jung Hotel, New Orleans, April 26–28, 1972.

PI MU EPSILON, Dartmouth College, Hanover, August 29–30, 1972.

SOCIETY FOR INDUSTRIAL AND APPLIED MATHEMATICS, Benjamin Franklin Hotel, Philadelphia, June 12–14, 1972 (20th Anniversary Celebration).

GEORGE BANTA COMPANY, INC., MENASHA, WISCONSIN

THE AMERICAN
MATHEMATICAL MONTHLY

(FOUNDED IN 1894 BY BENJAMIN F. FINKEL)

THE OFFICIAL JOURNAL OF

THE MATHEMATICAL ASSOCIATION OF AMERICA, INC.

VOLUME 78 NUMBER 9

PART I

CONTENTS

(*Continued on inside cover*)

NOVEMBER 1971

NOTICE TO AUTHORS

Specialized research is usually unsuitable; see *Statement of Policy* (vol. 76, p. 2). Manuscript preparation: Please use the *Manual for Monthly Authors* (vol. 78, p. 1) and follow the format in current issues of the MONTHLY. Manuscripts should be typewritten, triple-spaced with wide margins; submit two copies and keep one for protection against loss.

Backlog: Main Articles 7 months, Math. Notes 8 months, Research Problems 6 months, Classroom Notes 7 months, Math. Education 6 months.

EDITORIAL CORRESPONDENCE AND MAIN ARTICLES: to HARLEY FLANDERS, American Mathematical Monthly, Tel Aviv University, Ramat Aviv, Israel (see Notice, vol. 77, 1970, p. 555); NOTES, etc.: to the corresponding Associate Editor; ADVERTISING CORRESPONDENCE: to RAOUL HAILPERN, Mathematical Association of America, SUNY at Buffalo, Buffalo, N. Y. 14214; CHANGE OF ADDRESS and SUBSCRIPTIONS: to A. B. WILLCOX, Mathematical Association of America, 1225 Connecticut Ave., N.W., Washington, D. C. 20036.

Annual dues for members of the Association (including a subscription to the American Mathematical Monthly) are $12.50. For nonmembers the subscription price is $18.00.

PUBLISHED BY THE ASSOCIATION at Washington, D. C., and Menasha, Wisconsin during the months of January, February, March, April, May, June–July, August–September, October, November, December.

Second-class postage paid at Washington, D. C., and additional mailing offices. Acceptance for mailing at special rate of postage provided for in the Act of February 28, 1925, embodied in Paragraph 4, Section 538, P. L. and R., authorized April 1, 1926.

PRINTED IN THE UNITED STATES OF AMERICA

SOME RECENT RESULTS ON TOPOLOGICAL MANIFOLDS

REINHARD SCHULTZ, Purdue University

Although topological spaces exist in great variety and can exhibit strikingly unusual properties, the main concern of topology has generally been the study of spaces which are relatively well-behaved. One particularly interesting class of examples is given by those spaces which locally look like Euclidean n-space R^n. Explicitly, a Hausdorff space X is called a **topological n-manifold** (without boundary) if each point of X has an open neighborhood which is homeomorphic to an open subset in R^n. Since open sets in R^m and R^n are homeomorphic if and only if $m=n$, the integer n is a homeomorphism invariant of X and is called the **dimension** of X. In this paper all manifolds under consideration are assumed to be second countable.

Topological manifolds arise naturally in several different ways. For example, they are useful in the qualitative study of differential equations inaugurated by Poincaré (compare [1]). Topological manifolds are also a natural generalization of the mathematical systems studied in non-Euclidean and Riemannian geometry. Many interesting results on topological manifolds are generalizations of older theorems originally proved for these and similar mathematical systems.

During the nineteen sixties important advances in the study of topological manifolds yielded a great deal of information on their basic geometric structure. In particular, two long standing conjectures regarding topological manifolds were shown to be systematically false (see Section 4). One of the most useful results on topological manifolds of dimension $\neq 4, 5$—their description in terms of attaching handles—will be discussed in Section 5. This result allows one to take certain theorems which had previously been proved under additional structural assumptions and generalize them to topological manifolds with only minimal changes in the proofs.

I wish to thank R. Kirby for his detailed comments on an earlier version of this paper.

1. Classification of topological manifolds. Before beginning our discussion, it will be useful to generalize the definition of topological n-manifolds to include the possibility of a boundary. Let R^n_+ be the set of points in R^n whose last coordinate is nonnegative. Then a **topological n-manifold with boundary** is a Hausdorff space X each point of which has an open neighborhood homeomorphic to an open subset of R^n or R^n_+.

Of course, the set of all points having neighborhoods homeomorphic to open subsets of R^n is a topological n-manifold without boundary as previously defined. It is easy to see that the set of such points is open and dense in M; this subset is called the **interior of** M and written Int M. The complement of Int M is called the **boundary** of M and written ∂M; it follows that ∂M is a topological $(n-1)$-

Reinhard Schultz received his Chicago Ph.D. in 1968, under Richard K. Lashof. His main research interest is differential topology, and he is currently Assistant Professor at Purdue. *Editor.*

manifold without boundary. The following theorem of M. Brown [**9**] is extremely important in the study of manifolds with boundary:

THEOREM 1.1. (Collar Neighborhood Theorem) *Let M be a manifold with boundary. Then there is an open neighborhood V of ∂M which is homeomorphic to $\partial M \times [0, 1)$ such that $\partial M \subseteq V$ corresponds to $\partial M \times \{0\}$.*

One of the most immediate problems regarding topological manifolds is their classification up to homeomorphism. The techniques of point set topology suffice for the classification of one-dimensional manifolds; this was completed during the second decade of the twentieth century (see [**37**] or [**41**]). There are only four different homeomorphism types of connected one-dimensional manifolds: The open interval, the half-open interval, the closed interval, and the circle.

The study of two-dimensional manifolds is somewhat more difficult and requires a systematic investigation of polyhedra in the Euclidean plane (e.g., see [**29**], [**30**], or [**41**]). One of the earliest results was the Jordan Curve Theorem, first proved correctly by Veblen in 1905 [**59**]. This theorem was augmented by a result of Schoenflies [**48**], and we may combine the two theorems into the following single statement:

THEOREM 1.2. (Jordan-Schoenflies Theorem). *Let X be a subset of R^2 which is homeomorphic to a circle. Then $R^2 - X$ has two components, one bounded and one unbounded, and X is the point set-theoretic frontier of each component. The homeomorphism from the unit circle to X extends to a homeomorphism from the unit disk to the closure of the bounded component of $R^2 - X$.*

This theorem is the basic result needed for the following theorem of Radó [**45**]:

THEOREM 1.3. *Any (unbounded) topological two-dimensional manifold M may be triangulated; i.e., there is a countable locally finite covering $\{T_i\}$ of M by compact subspaces satisfying:*

(i) *There are canonical homeomorphisms h_i from T_i to the solid triangle*

$$\{(x, y) \in R^2 \mid x \geqq 0, y \geqq 0, \text{ and } x + y \leqq 1\}.$$

(ii) *Under these homeomorphisms any nonempty intersection $T_i \cap T_j$ corresponds to either a common side or a common vertex.*

The classification of two-dimensional manifolds up to homeomorphism then follows from a study of triangulated manifolds (e.g., see [**18**], [**30**], [**46**]). Results of Moise imply that the classification of three-dimensional manifolds reduces to a study of triangulated three-dimensional manifolds (e.g., [**35**], [**36**]), and a classification scheme in the compact case exists modulo Conjecture 1.4 below.

If X is any arcwise connected space, then X is said to be **simply connected**

if any continuous map from the unit circle in R^2 to X (i.e., a *closed curve* in X) extends to a continuous map of the unit disk. Given this definition, we may state the following conjecture made by Poincaré in 1904 [**44**]:

CONJECTURE 1.4 (Poincaré Conjecture). *Let M be a compact topological 3-manifold without boundary that is simply connected. Then M is homeomorphic to the unit sphere in R^4 (i.e., the 3 dimensional sphere).*

Relatively little is known about four-dimensional manifolds; the direct approach used in lower dimensions becomes increasingly complicated as the dimension increases, and in four dimensions the problems involved becomes forbiddingly difficult. There is a marked change, however, when one considers manifolds of dimension at least five. In this case one has enough space in which to make geometric constructions involving circles and disks almost at will. A particular consequence of this freedom of construction is that no general classification scheme for compact topological manifolds exists in any dimension $\geqq 5$ (compare [**5**, pp. 375–376]); for the freedom in constructing higher dimensional manifolds implies that any classification scheme would yield a solution to the word problem for finitely presented groups (see [**47**, Ch. XII] for a discussion of the latter problem).

2. Generalized Schoenflies and Poincaré Conjectures. The Jordan curve theorem was soon generalized to higher dimensions by Brouwer ([**7**]; also see [**14**, §18] or [**54**]). However, Antoine [**4**] and Alexander [**3**] constructed examples of subspaces X in R^3 that are homeomorphic to the unit sphere in R^3 but are not the frontiers of subspaces homeomorphic to the unit disk; counterexamples similar to Alexander's exist in all higher dimensions. On the other hand, Alexander also proved that X bounds a disk if it is a polyhedron in R^3 [**2**]. Around 1960 B. Mazur [**31**], M. Morse [**38**], and M. Brown [**8**] proved results implying the following generalization of the Jordan-Schoenflies theorem:

THEOREM 2.1. (Generalized Schoenflies Theorem). *Let X be a subset of R^n that is homeomorphic to the unit sphere, and assume that the closure of the bounded component of R^n-X is a topological n-manifold with boundary. Then the homeomorphism from the sphere to X extends to a homeomorphism from the disk to the closure of the bounded component of R^n-X.*

About the same time that the Generalized Schoenflies Theorem was proved, Smale [**53**], Stallings [**55**], and Zeeman [**64**] proved a generalization of the Poincaré Conjecture (1.4 above) in all dimensions greater than four; however, their proofs required additional structure on the manifolds under consideration (i.e., they had to be differential or combinatorial manifolds as defined in Section 3). Several years later Newman gave a proof of this result for topological manifolds using his generalization of Stallings' techniques and arguments of E. H. Connell [**42**]. For completeness, we state the result below:

THEOREM 2.2. (Generalized Poincaré Conjecture). *Let M be a compact topo-*

logical n-manifold ($n \geqq 5$) without boundary that is $\frac{1}{2}(n-1)$-connected if n is odd and $\frac{1}{2}n$-connected if n is even. Then M is homeomorphic to the unit sphere in R^{n+1}.

REMARKS 1. A topological space X is said to be **k-connected** if any continuous map from the unit sphere in R^{k-m+1} (for any $m \geqq 0$) extends to a map of the unit disk.

2. We already noted that the three-dimensional case of Theorem 2.2 is unknown; the four-dimensional case is also unknown.

The proof breaks down in dimensions 3 and 4 because in these cases there is not enough room in the manifold to make all the constructions needed in the proof (compare the last paragraph of Section 1).

Smale's proof of the generalized Poincaré Conjecture (most of whose details are independently due to A. H. Wallace [**62**]) was a central technique in the theory of surgery on manifolds developed by Kervaire, Milnor, S. P. Novikov, W. Browder, and C. T. C. Wall (for a definitive account see Wall's book [**61**]). Wall's theory in turn was important in studying the following elaboration of the Generalized Schoenflies Conjecture:

CONJECTURE 2.3. (Annulus Conjecture). *Let $A \subseteq R^{n+1}$ be a compact topological $(n+1)$-manifold whose boundary is homeomorphic to a disjoint union of two copies of the unit sphere in R^{n+1}. Then A is homeomorphic to the closed annulus in R^{n+1} bounded by the spheres of radius* 1 *and* 2.

If this conjecture were false for $n=1$ or 2, then an argument of Brown and Gluck [**10**, p. 42] would imply that the compact unbounded topological manifold A' formed from A by gluing together the two components of the boundary of A could not be triangulated. Hence the conjecture is certainly true in these dimensions by *reductio ad absurdum* (more elementary arguments are also possible). In [**19**] Kirby gave an elegant argument which reduced the proof of the annulus conjecture for $n \geqq 4$ to a problem which could be handled by means of Wall's surgery theory. This surgery theoretical problem was solved independently by Wall [**60**] and W.-C. Hsiang and Shaneson ([**15**], [**16**]); thus Conjecture 2.3 is true except possibly in the case $n=3$.

3. Differentiable and Combinatorial Manifolds. In this section we shall describe the kinds of "additional structure" often associated to topological manifolds and mentioned in the previous sections.

The topological manifolds appearing in analysis and differential geometry usually satisfy the conditions appearing in the following definition:

DEFINITION. A topological n-manifold is **smoothable** if there is a collection of pairs $\{(U_\alpha, h_\alpha)\}_{\alpha \in A}$ satisfying:

(i) U_α is an open subset in R^n.

(ii) The map $h: U_\alpha \to M$ is a homeomorphism onto an open subset.

(iii) The functions $h_\beta^{-1}h_\alpha : h_\alpha^{-1}h_\beta(U_\beta) \to h_\beta^{-1}h_\alpha(U_\alpha)$ are functions of class C^r for some $r \geqq 1$.

If U and W are open subsets of Euclidean spaces, recall that a map $f: U \rightarrow W$ is a **function of class** C^r if the coordinate functions f^i defined by $f(w) = (f^1(w), \cdots, f^m(w))$ each have all possible partial derivatives of order r and these functions are continuous; a function is C^∞ if it is C^r for every positive integer r. Two collections $\{(U_\alpha, h_\alpha)\}$ and $\{(V_\beta, k_\beta)\}$ satisfying (i)–(iii) are **equivalent** if their union satisfies property (iii); it follows that every collection $\{(U_\alpha, h_\alpha)\}$ is equivalent to a unique maximal collection $\mathfrak{A}$ which is called a **smooth atlas** for M **of class** C^r. A **differential** (or **smooth**) n-**manifold** is a pair $(M, \mathfrak{A})$ consisting of a smoothable n-manifold M and a smooth atlas $\mathfrak{A}$. We shall always assume that the atlas is smooth of class C^∞, since it is known that any C^r atlas corresponds to a unique C^∞ atlas [**40**, Sections 4 and 5].

More generally, if Γ is any reasonable family of continuous functions from open sets in R^n to open sets in R^n (technically a **pseudogroup**; see [**22**]), then it is possible to define a Γ atlas and a Γ n-manifold. In topological investigations Γ is usually taken to be the C^r functions defined above or the piecewise linear (PL) functions defined below. Thus in order to define a piecewise linear n-manifold, it is only necessary to specify which mappings on open subsets of Euclidean space are piecewise linear; this requires a succession of definitions.

DEFINITION. Let $x_0, \cdots, x_n$ be points in R^m such that $x_1 - x_0, \cdots, x_n - x_0$ are linearly independent. Then the n-dimensional **simplex** (or n-simplex) with vertices $x_0, \cdots, x_n$ is the set of all linear combinations $y = \sum t_i x_i$, where each t_i is nonnegative and $\sum t_i = 1$ (the last condition and linear independence imply that the t_i are unique). The x_i are called the **vertices** of the simplex.

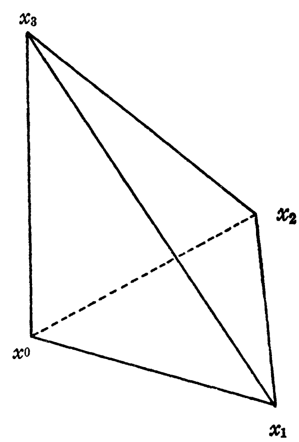


FIG. 1

A simplex is actually a generalized version of a triangle. It is immediate from the definition that a 1-simplex is a line segment and a 2-simplex is a solid triangle. Furthermore, a 3-simplex is a tetrahedron (see Figure 1).

DEFINITION. Let A be a simplex with vertices $a_i (1 \leqq i \leqq n)$, and let V be any

real vector space. A function $f: A \to V$ is **affine linear** provided $y \in A$ and $y = \sum t_i a_i$ with $\sum t_i = 1$ imply $f(y) = \sum t_i f(a_i)$.

DEFINITION. Let U and V be any subsets of R^n. A continuous function $f: U \to V$ is **piecewise linear** (or **PL**) if there is a countable locally finite covering $\mathfrak{B}$ of U by simplexes such that f is an affine linear map on each element of $\mathfrak{B}$.

REMARK. Any open subset of R^n_+ has many countable locally finite coverings by simplexes.

EXAMPLES 1. Let $f: R^n \to R^n$ be an affine transformation; i.e., $f(x) = Lx + y$, where L is a linear transformation. Then f is automatically affine linear on every simplex in R^n (compare [**6**, p. 272]).

2. Let $f: R^2 \to R^2$ be given by $f(x, y) = (x, y)$ if $y \geqq 0$ and $(x, 2y)$ if $y \leqq 0$. Then f is affine linear on any simplex contained in either the upper or lower half plane.

3. Let f be the map which sends the solid regular pentagon $ABCDE$ to the solid irregular pentagon $A'B'C'D'E'$ in Figure 2 by stretching the triangle OXY into $O'X'Y'$.

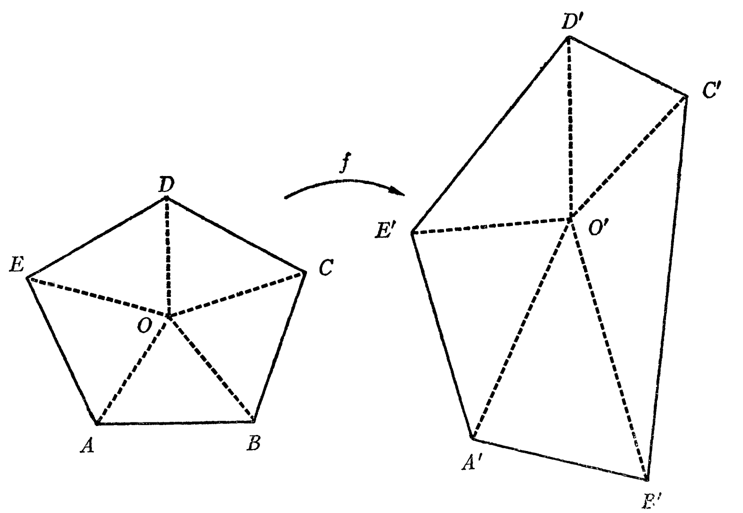


FIG. 2

A fundamental theorem of Cairns and Whitehead states that any smooth manifold determines a basically unique PL manifold ([**11**], [**63**], [**40**, Pt. II]). However, a PL manifold need not be determined by a smooth manifold (a result of Thom [**58**]), and two distinct smooth manifolds may determine the same PL manifold (a result of Milnor [**32**]). A comprehensive study of the relationship between smooth and PL manifolds appears in [**25**].

In the following section we shall discuss the parallel problem regarding the existence and uniqueness of PL manifolds associated to a given topological manifold.

REMARK. For historical reasons the study of PL manifolds and related objects is frequently called **combinatorial topology** and PL manifolds are often called **combinatorial manifolds.**

4. The Triangulation Conjecture and the Hauptvermutung. The following conjectures were formulated (in roughly equivalent form) soon after the establishment of combinatorial topology as a subject in its own right.

TRIANGULATION CONJECTURE. *Any topological n-manifold has a* PL *atlas.*

HAUPTVERMUTUNG FOR MANIFOLDS. *Any two homeomorphic* PL *n-manifolds are equivalent as* PL *manifolds.*

The results quoted in Section 1 imply that the first conjecture is true if $n \leqq 3$. Similarly, the second conjecture is true if $n \leqq 3$ ($n = 1$, straightforward; $n = 2$, see Papakyriakopoulos [**43**]; $n = 3$, see Moise [**35**], [**36**]). The solution of the generalized Poincaré conjecture in higher dimensions implies that the second conjecture is true for PL manifolds homeomorphic to spheres of dimension at least five. A fairly strong version of the *Hauptvermutung* for simply connected manifolds was proved by Lashof and Rothenberg [**26**], and Sullivan ([**56**], [**57**]); in the next paragraph we shall discuss subsequent results which eliminated the simple connectivity assumption (see Theorem 4.2).

Kirby's reduction of the Annulus Conjecture, other results appearing in [**19**], and consequences of these results due to Lees [**28**] led directly to initial results on the Triangulation Conjecture due to Lashof [**23**]. These theorems and computations of Casson, Wall, Hsiang, and Shaneson ([**15**], [**16**], [**60**]) in turn led to the following strong results on the Triangulation Conjecture and the *Hauptvermutung* due to Lashof and Rothenberg ([**27**], [**24**]), and Kirby and Siebenmann [**20**]:

THEOREM 4.1. *Let M be a topological manifold of dimension at least six* (*or five in the unbounded case*), *and assume that the four-dimensional cohomology group $H^4(M; Z_2)$ is zero. Then M has a* PL *atlas.*

THEOREM 4.2. *Let M be a* PL *manifold satisfying the above dimensional restriction, and assume that the three-dimensional cohomology group $H^3(M; Z_2)$ is zero. Then any* PL *manifold homeomorphic to M is equivalent to M as a* PL *manifold.*

REMARK. For the sake of completeness we shall describe the cohomology groups $H^k(M; Z_2)$ in a geometric manner exploited by Sullivan in his proof of the earlier version of Theorem 4.2; for a more standard description of $H^k(M; Z_2)$ see [**14**, §23] or any algebraic topology text. If X is any topological space, a **smooth k-manifold in X** is a continuous function $f: V \rightarrow X$, where V is a compact smooth k-dimensional manifold. An element in $H^k(X; Z_2)$ is then a function which assigns to each k-manifold in X an element of Z_2 subject to certain consistency conditions which are straightforward but a little too technical to de-

scribe here (see [**12**, §8] or [**56**] for further discussion; the description does not generalize to odd primes).

The restrictions on cohomology appearing in the above theorems were also shown to be unnecessary if $H^3(\mathrm{Top/PL}; Z_2)=0$, where Top/PL is a topological space arising from the geometry of the proof of 4.1 and 4.2. However, Siebenmann (first alone and later jointly with Kirby) constructed examples which implied that $H^3(\mathrm{Top/PL}; Z_2)$ is nonzero; it followed quickly that both the Triangulation Conjecture and the *Hauptvermutung* were systematically false in every dimension greater than four.

There are very simple manifolds which yield contradictions to the *Hauptvermutung*. For example, consider the cartesian product $S^3\times T^2$ of the unit sphere in R^4 with the two-dimensional torus T^2. This product is a smooth manifold and consequently determines a unique PL manifold; results of Shaneson combined with $H^3(\mathrm{Top/PL}; Z_2)\neq 0$ imply the existence of a PL 5-manifold M^5 which is homeomorphic to $S^3\times T^2$ but inequivalent to $S^3\times T^2$ as a PL manifold ([**49**], [**50**]).

5. Handlebody theory for topological manifolds. In one sense the Kirby-Siebenmann results are disappointing because they disprove two conjectures which would have reduced the study of topological manifolds to combinatorial topology. On the other hand, the results used in the proof of 4.1 and 4.2 yield a convenient method for decomposing topological manifolds of dimension at least six, which will be discussed in this section.

Throughout this section S^p will denote the unit sphere in R^{p+1} and D^{p+1} will denote the unit disk in R^{p+1}. It follows from the definitions that D^{p+1} is a topological $(p+1)$-manifold with boundary, and its boundary is S^p.

DEFINITION. Let V be a topological n-manifold with boundary, and let $f: S^{k-1}\times D^n\to\partial V$ be a one-to-one continuous mapping. Then the manifold W obtained by **attaching a k-handle to V along** f is the disjoint union of V and $D^k\times D^{n-k}$ modulo the identification of $f(S^{k-1}\times D^{n-k})\subseteq V$, with $S^{k-1}\times D^{n-k}\subseteq D^k\times D^{n-k}$ (see Figure 3 for an illustration).

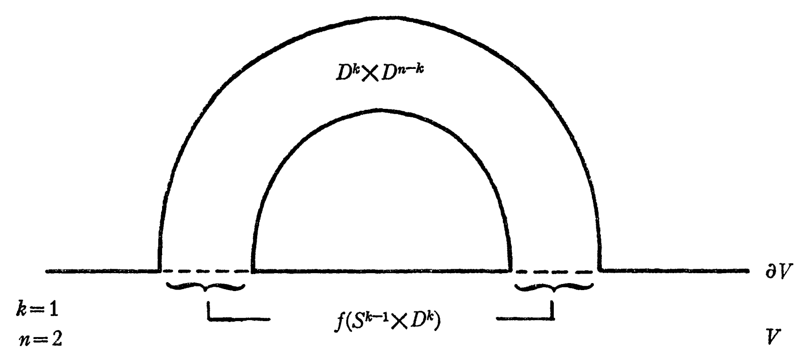


FIG. 3

This construction dates back to the beginnings of the study of manifolds. For example, the classification theorem for compact orientable 2-manifolds may be written as follows (compare [**30**]):

THEOREM 5.1. *Let M be a compact orientable 2-manifold. Then $M=\partial Q$, where Q is formed by attaching 1-handles to D^3.*

The work of Marston Morse on critical point theory implies that any smooth manifold M may be constructed by successively attaching handles (e.g., see [**33**], [**34**], [**39**]); in terms of the definition below, M has a *handle decomposition.* Standard results of combinatorial topology imply a similar result for PL manifolds [**17**, p. 226].

In the definition below, $X\cup\partial X\times[0, 1]$ will be interpreted to mean the disjoint union modulo the identification of $y\in\partial X$ with $(y, 0)\in\partial X\times\{0\}$. We shall assume the manifold M discussed below is either unbounded or compact in order to simplify the definition.

DEFINITION. Let M be a topological manifold. A **handle decomposition** of M is a (finite or denumerable) sequence of compact subspaces $\{M_j\}_{j\in J}$ (J a well-ordered subset of the integers) satisfying:

(i) $M=\bigcup_{j\in J} M_j$ and each M_j is a compact manifold with boundary.

(ii) For all $j\in J$ we have $M_j\subseteq\operatorname{Int} M_{j+1}$; in fact, M_{j+1} is formed by attaching a k-handle to $M_j\cup\partial M_j\times[0, 1]$ (provided j is not maximal in J).

The following result of Kirby and Siebenmann is a straightforward consequence of the arguments used to prove 4.1 and 4.2 [**21**]:

THEOREM 5.2. *Any topological manifold of dimension greater than five has a handle decomposition.*

This is one case of a general principle implicit in [**21**]; namely, results which work for smooth and PL manifolds in dimensions greater than five also work for topological manifolds in dimensions greater than five. Some particular examples are the theorems of Siebenmann [**51**] and Farrell [**13**] and the surgery theory presented in [**61**].

Since any topological manifold of dimension $\leqq 3$ has a PL atlas, and hence a handle decomposition, the only unknown cases occur in dimensions 4 and 5. The nonvanishing of $H^3(\mathrm{Top/PL}; Z_2)$ implies the following negative result due to Siebenmann [**52**]:

THEOREM 5.3. *For $n=4$ or 5 (possibly both) there exists a compact unbounded topological n-manifold that has no handle decomposition.*

Siebenmann also proves in [**52**] that certain fundamental theorems on smooth and PL manifolds in dimensions greater than five fail somewhere in dimensions three, four, and five. Precise knowledge of where these failures occur would be a useful addition to our relatively meager knowledge of manifolds in these dimensions.

References

1. R. Abraham, Foundations of Mechanics, Benjamin, New York, 1967.

2. J. W. Alexander, On the subdivision of 3-space by a polyhedron, Proc. Nat. Acad. Sci. U.S.A., 10 (1924) 6–8.

3. ———, An example of a simply connected surface bounding a region which is not simply connected, Proc. Nat. Acad. Sci. U.S.A., 10 (1924) 8–10.

4. L. Antoine, Sur l'homéomorphie de deux figures et de leurs voisinages, J. Math. Pures Appl., (8) 4 (1921) 221–325.

5. D. Barden, Simply connected five-manifolds, Ann. of Math., 82 (1965) 365–385.

6. G. Birkhoff and S. MacLane, Survey of Modern Algebra, Third edition, Macmillan, New York, 1965.

7. L. E. J. Brouwer, Beweis des Jordanschen Satzes für den n-dimensionalen Raum, Math. Ann., 71 (1912) 314–319.

8. M. Brown, A proof of the generalized Schoenflies Theorem, Bull. Amer. Math. Soc., 66 (1960) 74–76.

9. ———, Locally flat embeddings of topological manifolds, Ann. of Math. 75 (1962), 331–341.

10. ———, and H. Gluck, Stable structures on manifolds I, II, III, Ann. of Math., 79 (1964) 1–58.

11. S. S. Cairns, Triangulation of the manifold of class one, Bull. Amer. Math. Soc., 41 (1935) 549–552.

12. P. E. Conner and E. E. Floyd, Differentiable Periodic Maps, Ergebnisse der Mathematik und ihrer Grenzgebiete, N. F. Bd. 33. Springer-Verlag, Berlin-Göttingen-Heidelberg, 1964.

13. F. T. Farrell, The obstruction to fibering a manifold over a circle, Bull. Amer. Math. Soc., 73 (1967) 737–740.

14. M. Greenberg, Lectures on Algebraic Topology, Benjamin, New York, 1967.

15. W. C. Hsiang and J. Shaneson, Fake tori, the annulus conjecture, and the conjectures of Kirby, Proc. Nat. Acad. Sci. U.S.A., 63 (1969) 687–691.

16. ———, Fake tori, Topology of Manifolds (Proc. Univ. of Georgia Institute, 1969), 18–51. Markham, Chicago, 1970.

17. J. F. P. Hudson, Piecewise Linear Topology, Benjamin, New York, 1969.

18. B. von Kérékjárto, Vorlesungen über Topologie, Die Grundlehren der mathematischen Wissenschaften, Bd. 1. J. Springer, Berlin, 1923.

19. R. C. Kirby, Stable homeomorphisms and the annulus conjecture, Ann. of Math., 89 (1969) 575–582.

20. ———, and L. C. Siebenmann, On the triangulation of manifolds and the Hauptvermutung, Bull. Amer. Math. Soc., 75 (1969) 742–749.

21. ———, Foundations of Topology, Notices, Amer. Math. Soc., 16 (1969) 848.

22. S. Kobayashi and K. Nomizu, Foundations of Differential Geometry, Vol. I., Interscience Tracts in Pure and Applied Mathematics, vol. 15. Interscience, New York, 1963.

23. R. Lashof, Lees' immersion theorem and the triangulation of manifolds, Bull. Amer. Math. Soc., 75 (1969) 535–538.

24. ———, The immersion approach to triangulation, Topology of Manifolds (Proc. Univ. of Georgia Institute, 1969), 52–56. Markham, Chicago, 1970.

25. ———, and M. Rothenberg, Microbundles and smoothings, Topology, 3 (1964/65) 357–388.

26. ———, Hauptvermutung for manifolds, Conference on the Topology of Manifolds (Michigan State, 1967), 81–105. Prindle, Weber, and Schmidt, Boston, 1968.

27. ———, Triangulation of manifolds I, II, Bull. Amer. Math. Soc., 75 (1965) 750–757.

28. J. Lees, Immersions and surgeries of topological manifolds, Bull. Amer. Math. Soc., 75 (1969) 529–534.

29. S. Lefschetz, Introduction to Topology, Princeton Mathematical Series, vol. 11. Princeton University Press, Princeton, 1949.

30. W. S. Massey, Algebraic Topology; an Introduction, Harcourt, Brace, and World, New York, 1967.

31. B. Mazur, On embeddings of spheres, Acta Math., 105 (1961) 1–17.

32. J. Milnor, On manifolds homeomorphic to the 7-sphere, Ann. of Math., 64 (1956) 399–405.

33. J. Milnor, Morse Theory (notes by M. Spivak and R. Wells), Annals of Mathematics Studies No. 51, Princeton University Press, Princeton, 1963.

34. ———, Lectures on the h-cobordism theorem (notes by L. Siebenmann and J. Sondow), Princeton University Press, Princeton, 1965.

35. E. E. Moise, Affine structures in 3-manifolds, V. The triangulation theorem, Ann. of Math., 56 (1952) 96–114.

36. ———, Affine structures in 3-manifolds, VIII. Invariance of knot types; local tame embedding, Ann. of Math., 59 (1954) 159–170.

37. R. L. Moore, Concerning simple continuous curves, Trans. Amer. Math. Soc., 21 (1920) 333–347.

38. M. Morse, A reduction of the Schoenflies extension problem, Bull. Amer. Math. Soc., 66 (1960) 113–115.

39. ———, and S. S. Cairns, Critical Point Theory in Global Analysis and Differential Topology, Pure and Applied Mathematics Series, Vol. 33, Academic Press, New York, 1969.

40. J. Munkres, Elementary Differential Topology (revised edition), Annals of Mathematics Studies No. 54, Princeton University Press, Princeton, 1966.

41. M. H.A. Newman, Elements of the Topology of Plane Sets of Points, Cambridge University Press, Cambridge, 1951.

42. ———, The engulfing theorem for topological manifolds, Ann. of Math., 84 (1966) 555–571.

43. C. D. Papakyriakopoulos, A new proof for the invariance of the homology groups of a complex (Greek), Bull. Soc. Math. Grèce, 22 (1943) 1–154.

44. H. Poincaré, Cinquième complément à l'Analysis situs, Rend. Circ. Mat. Palermo, 8 (1904) 45–110.

45. T. Radó, Über den Begriff von Riemannsche Fläche, Acta Sci. Math. (Szeged), 2 (1924/26) 101–120.

46. I. Richards, On the classification of noncompact surfaces, Trans. Amer. Math. Soc., 106 (1963) 259–269.

47. J. Rotman, Theory of Groups: an Introduction, Allyn and Bacon, Boston, 1965.

48. A. Schoenflies, Beiträge zur Theorie der Punktmengen III, Math. Ann., 62 (1906) 286–328.

49. J. Shaneson, Wall's surgery obstruction groups for $G\times Z$, Ann. of Math., 90 (1969) 296–334.

50. ———, Non-simply connected surgery and some results in low-dimensional topology, Comment. Math. Helv., 45 (1970) 333–352.

51. L. C. Siebenmann, The obstruction to finding a boundary for an open manifold of dimension greater than five, Ph.D. Thesis, Princeton University, 1965.

52. ———, Disruption of low-dimensional handlebody theory by Rohlin's theorem, Topology of Manifolds (Proc. Univ. of Georgia Institute, 1969), 57–76. Markham, Chicago, 1970.

53. S. Smale, Generalized Poincaré's conjecture in dimensions greater than four, Ann. of Math., 74 (1961) 391–406.

54. E. H. Spanier, Algebraic Topology, McGraw-Hill, New York, 1966.

55. J. R. Stallings, Polyhedral homotopy-spheres, Bull. Amer. Math. Soc., 66 (1960) 485–488.

56. D. Sullivan, On the Hauptvermutung for manifolds, Bull. Amer. Math. Soc., 73 (1967) 598–600.

57. ———, Triangulating and smoothing homotopy equivalences and homeomorphisms, Geometric Topology Seminar Notes, mimeographed, Princeton University, 1967.

58. R. Thom, Les classes caractéristiques de Pontrjagin des variétés triangulées, Symposium internacional de topología algebraica (Mexico, 1956), 54–67. Universidad Nacional Autónoma de México y UNESCO, Mexico, 1958.

59. O. Veblen, Theory of plane curves in nonmetrical analysis situs, Trans. Amer. Math. Soc., 6 (1905) 83–98.

60. C. T. C. Wall, On homotopy tori and the annulus theorem I, Bull. London. Math. Soc., 1 (1969) 95–97.

61. ———, Surgery on Compact Manifolds, London Mathematical Society Monographs No. 1, Academic Press, New York, 1970.

62 A. H. Wallace, A geometric method in differential topology, Bull. Amer. Math. Soc., 68 (1962) 533–542.

63. J. H. C. Whitehead, On C^1 complexes, Ann. of Math., 41 (1940) 809–824.

64. E. C. Zeeman, The Poincaré conjecture for $n \geqq 5$, Topology of 3-manifolds and Related Topics (Proc. Univ. of Georgia Institute, 1961), 198–204. Prentice-Hall, Englewood Cliffs, N. J., 1962.

REMINISCENCES OF AN OCTOGENARIAN MATHEMATICIAN

L. J. MORDELL,[1] St. John's College, Cambridge, England

It is customary for the fellows of St. John's College, Cambridge, to dine privately on December 27, the birthday of St. John, the Evangelist. The Master proposes a toast to those fellows who have attained the age of eighty since the preceding December 27, and asks each of them to give a talk. As I became eighty on January 28, 1968, it was my turn to do so.

I started off by saying that this was a really great occasion in my life and that I was very grateful to our College for making it possible. I said that it was not an easy matter to make an appropriate speech on such an occasion. Fortunately it was not too difficult for me to do so, as I have recently been reading a book by the well-known and popular American author Dale Carnegie, entitled *How to Stop Worrying and Start Living.* In this, he makes the cogent remark that no man is so happy as when he is talking about himself. He says nothing about the feelings of his listeners.

There are two reasons why I propose to make myself thoroughly and unashamedly happy by talking about myself. The first is that on several occasions, both in England and America, I have been told that I am a legendary character. As it occurs to me that most legendary characters, for example King Arthur, are dead, I wish to show that I have actually existed and am very much alive, and so I shall give some account of the subject so that there will be no doubt about the matter.

The second reason is that there have been many stories, mostly apocryphal, as to how I, a natural born American, came to study at St. John's College. The reason is a very simple and natural one. I do not mean to be boastful or vainglorious, and I wish to apologize if I seem so and to crave your indulgence.

[1] This talk was presented to the Philadelphia Section of the MAA on Nov. 22, 1969 at Swarthmore College. It was given in part to the Fellows of St. John's College on Dec. 27, 1968 and again to the Adam Society, St. John's College, on March 5, 1969.

you nor I nor anybody else knows what makes a mathematician tick. It is not a question of cleverness. As I have already said, I know many mathematicians who are far abler and cleverer than I am, but they have not been so lucky. An illustration may be given by considering two miners. One may be an expert geologist, but he does not find the golden nuggets that the ignorant miner does.

In some ways, a mathematician is not responsible for his activities. One sometimes feels there is an inner self occasionally communicating with the outer man. This view is supported by the statements made by H. Poincaré and J. Hadamard about their researches. I remember once walking down St. Andrews Street some three weeks after writing a paper. Though I had never given the matter any thought since then, it suddenly occurred to me that a point in my proof needed looking into.

I am very grateful to my inner self for his valuable help in the solution of some important and difficult problems that I could not have done otherwise.

I commenced this talk by saying a toast had been drunk to me by the Master and Fellows of St. John's College. I might conclude by reciting one sent to me by Professor L. Moser. Of him, it was said that he was writing a book and taking so long about it that his publishers became very much worried and went to see him. He said he was very sorry about the delay, but he was afraid that the book might have to be a posthumous one. Well, he was told, please hurry up with it.

Moser's toast was as follows:

Here's a toast to L. J. Mordell,
young in spirit, most active as well,
He'll never grow weary,
of his love, number theory,
The results he obtains are just swell.

ALGEBRAIC CHARACTERIZATION OF SOME CLASSICAL COMBINATORIAL PROBLEMS

E. T. ORDMAN, University of Kentucky

1. Introduction. The numbers a_n,

$$a_n = \frac{1}{n}\binom{2n-2}{n-1}, \tag{1}$$

known as the **Catalan numbers,** occur in a wide variety of combinatorial problems. For example, a_n is the number of elements in the sets A_n, E_n, S_n, where:

Professor Ordman received his A.B. from Kenyon College, and his Princeton Ph.D. in 1969, having worked under John Stallings at the Univ. of Calif., Berkeley during 1967–69, supported by a Danforth Fellowship. Since 1969 he has been Assistant Professor at the Univ. of Kentucky. His main research is in group theory problems motivated by topology. *Editor.*

A_n is the collection of noncommutative nonassociative binary products of a single generator taken n times. For instance,

$$A_4 = \{a(a(aa)),\ a((aa)a),\ (aa)(aa),\ (a(aa))a,\ ((aa)a)a\},$$

so $a_4 = 5$.

E_n is the collection of ways in which a fixed regular $(n+1)$-gon in the plane may be divided into triangles by $n-2$ diagonals which do not intersect in its interior.

S_n is the collection of sequences of $2n-2$ terms $(x_1, x_2, \cdots, x_{2n-2})$, where each $x_1 = \pm 1$, subject to the conditions $x_1+x_2+\cdots+x_{2n-2}=0$ and $x_1+x_2+\cdots+x_k \geqq 0$ for $1 \leqq k \leqq 2n-2$.

We call a set $Q=Q_1 \cup Q_2 \cup Q_3 \cdots$ of **Catalan type** if formula (1) gives the number of elements of Q_n. Proofs that various sets (including the above) are of Catalan type appear for instance in [**1**], [**4**], [**5**, Problem 7], [**6**], and [**10**, Problems 54, 83, 84]. Our object is to study the algebraic structure underlying sets of Catalan type, in order to help recognize such sets and answer some related questions. We shall not reproduce a proof that formula (1) gives the number of elements of A_n, but we shall prove that A_n, E_n, and S_n may be placed in one-one correspondence.

1.2. Published proofs that a given set Q is of Catalan type sometimes set up a one-one map between Q_n and A_n or some other set known to have the desired number of elements; more often they establish that Q_1 has one element and establish the recurrence (2)

$$\sum q_k q_{n-k} = q_n \qquad (\text{sum for } 1 \leqq k \leqq n-1), \tag{2}$$

where Q_n has q_n elements. Since $a_1=1$ and the a_i satisfy this recurrence, $q_n=a_n$ by induction. Recurrence (2) is most often established by providing a way of "factoring" an element of Q_n uniquely as an element of Q_k and an element of Q_{n-k}, for some k. Thus the set Q is provided with a multiplication. For instance, multiplication in $A=A_1 \cup A_2 \cup \cdots$ is given by $(b, c) \to (b)(c)$ for every b and c in A (the parentheses around b or c being omitted if it is in A_1). In the future, we shall write the product of b and c simply as bc, when no confusion will result.

1.3. The sets A, $E=E_1 \cup E_2 \cup \cdots$, and $S=S_1 \cup S_2 \cup \cdots$ have various actions on them. We give some examples to be discussed later:

The operation "Mirror image" is a map $M: A \to A$. For instance M interchanges $(a(aa))a$ and $a((aa)a)$; but note $M((aa))aa)) = (aa)(aa)$. Clearly $M^2 = I$, the identity map on A.

The operation "reflection in a vertical line" is a map $\hat{M}: E \to E$. Like M, it is of order 2. The operation "rotation counterclockwise through $2\pi/(n+1)$ radians" is a map $\rho_n: E_n \to E_n$, with the property that ρ_n^{n+1} is the identity map of E_n.

S has a conspicuous map $\lambda: S \to S$ of order 2, given by

$$\lambda(x_1, x_2, \cdots, x_{2n-2}) = (-x_{2n-2}, -x_{2n-3}, \cdots, -x_1).$$

For instance, $\lambda(+1, +1, -1, -1, +1, -1) = (+1, -1, +1, +1, -1, -1)$.

1.4. *Problems*. Two of the maps of 1.3, as well as other "reasonable" maps we shall not study, fail to have various desirable properties. For instance, for none of the published "multiplications" $f: E_n \times E_m \to E_{n+m}$ does a relation such as

$$\rho_{n+m} f(b, c) = f(\rho_n b, \rho_m c), \tag{3}$$

for all b in E_n and c in E_m hold. Might f be defined to make (3) hold?

It is easy to see that no one-one onto map between A_3 and S_3 can carry M to λ; for M interchanges $(aa)a$ with $a(aa)$, while λ preserves both $(+1, +1, -1, -1)$ and $(+1, -1, +1, -1)$. Is there a "natural" map $M_S: S \to S$ corresponding to M? How does it relate to λ?

1.5. Section 2 is a discussion of the algebraic structure A, characterizing the structure itself, its automorphisms, and isomorphisms between it and other structures. Section 3 applies these methods to E, solving the problem about ρ_n stated above (Theorem 3.3). Section 4 contains a discussion of S, in particular the relation between M and λ.

1.6. *History*. The sets A_n, E_n, S_n, and many others like them, have long been well known. Dörrie [**5**, Problem 7] traces E_n back to a 1751 problem of Euler and A_n to an 1838 paper of Catalan. By 1859 Cayley [**4**] observed the connection between A_n and the problem of enumerating certain graphs. A wide variety of graph problems are connected with these numbers; similar numbers appear for instance in Tutte [**9**]. An extensive bibliography appears in Brown [**3**].

The problem of the number of elements in S_n is a special case of the ballot problem. Elementary discussions of related problems occur in [**6**] and [**10**, Problems 54, 83, 84]; a longer discussion and bibliography which includes references to a number of equivalent problems appears in Takács [**8**].

A number of other structures have also been put on the set A; see for instance [**7**], in which each A_n is made into a lattice.

2. Characterizations of the operation with no relations. Let us regard A as a set with a binary operation and a generator a. We observe first that the only automorphism of A is trivial.

2.1. THEOREM. *Let $f: A \to A$ satisfy $f(bc) = f(b)f(c)$, for all b and c in A. Then the following are equivalent*:

(a) *f is onto,*

(b) $f(a) = a$,

(c) *f is the identity map.*

Proof. If f is onto, there is some b in A with $f(b) = a$. If $b \neq a$, then $b = cd$ for some c and d in A; hence $a = f(cd) = f(c)f(d)$. But since a cannot be factored, this is impossible, so $b = a$ as desired. Suppose next that $f(a) = a$. Clearly $f(aa) = f(a)f(a) = aa$; also $f((aa)a) = f(aa)f(a) = (aa)a$, and by induction $f(bc) = f(b)f(c) = bc$, so f is the identity map $I: A \to A$. Finally, I is clearly onto.

In 1.3 we defined the "mirror image" map $M: A \to A$. It is clear that $M(a) = a$ and $M(bc) = M(c)M(b)$ for all b and c in A.

2.2. THEOREM. *Let $F:A\to A$ satisfy $f(bc)=f(c)f(b)$. Then the following are equivalent*:

(a) *f is onto*,
(c) $f(a)=a$,
(d) $f=M$.

Proof. Similar to 2.1.

M is the unique *anti-automorphism* of A. Clearly, no other map arising later in the discussion can be an automorphism or anti-automorphism.

A set Q with a binary operation $*$ is **isomorphic** to A if there is a one-one onto map f from Q to A such that $f(b*c)=f(b)f(c)$ for all b and c in Q. Then f is an **isomorphism**; a one-one onto map g such that $g(b*c)=g(c)g(b)$ will be called an **anti-isomorphism.**

2.3. COROLLARY. *Suppose $(Q, *)$ is isomorphic to A. Then there is a unique isomorphism $f:Q\to A$ and a unique anti-isomorphism $Mf:Q\to A$.*

Proof. If $f_1:Q\to A$ and $f_2:Q\to A$ are isomorphisms, $f_1f_2^{-1}$ and $f_2f_1^{-1}$ are automorphisms of A. Hence $f_1f_2^{-1}=I=f_2f_1^{-1}$ and $f_2=f_1$. Similarly if g_1 and g_2 are anti-isomorphisms, $g_1g_2^{-1}=I=g_2g_1^{-1}$ and $g_1=g_2$. If an isomorphism f exists, then $Mf(bc)=M(f(b)f(c))=Mf(c)Mf(b)$ so Mf is the unique anti-isomorphism.

To make it easier to recognize structures isomorphic to A, we introduce some terminology. If Q is a set and f is an operation $f:Q\times Q\to Q$, the structure (Q, f) will be called **graded** if $Q=Q_1\cup Q_2\cup\cdots$ and $f(Q_n\times Q_m)\subset Q_{n+m}$ for all $n, m\geqq 1$. If Q and R are graded sets, a map $g:Q\to R$ will be called **level-preserving** if $g(Q_n)\subset R_n$ for all n.

If (Q,f) is graded, Q_1 is not in the image $f(Q\times Q)$. We say **factoring is possible** if $f:Q\times Q\to Q\backslash Q_1$ is onto. We say **factoring is unique** if $f:Q\times Q\to Q\backslash Q_1$ is both onto and one-one; that is, if each q not in Q_1 has a unique expression $q=f(b, c)$. In particular, factoring in A is unique.

2.4. THEOREM. *Let $(Q, *)$ be a graded structure in which factoring is unique. Suppose Q_1 has exactly one element. Then there is a level-preserving isomorphism from $(Q, *)$ to A.*

Proof. The unique map $f:Q_1\to A_1$ is one-one onto. If f has been defined to be one-one, onto, level-preserving, and to preserve multiplication (when possible) from $Q_1\cup Q_2\cup\cdots\cup Q_n$ to $A_1\cup\cdots\cup A_n$, define f on Q_{n+1} by $f(q)=f(b)f(c)$, where $q=b*c$. This is well-defined since Q has unique factorization, and $f(q)$ is in A_{n+1} since f is level-preserving on b and c by the induction hypothesis. Finally $f:Q_{n+1}\to A_{n+1}$ is one-one onto; for if d is in A_{n+1}, then d can be uniquely written $d=uv$, and by the induction hypothesis there are unique b and c with $f(b)=u$ and $f(c)=v$. Now $b*c$ is in Q_{n+1} and is the unique element of Q_{n+1} mapping to $uv=d$.

It would be possible to change the hypotheses of 2.4 so that $(Q, *)$ is not graded, but $*:Q\times Q\to Q\backslash\{a\}$ is one-one, for some a in Q. One would then have

to require that a "generate" Q, to avoid the possibility of elements of "infinite length."

2.5. COROLLARY. *Let $(Q, *)$ be a graded structure isomorphic to A. Then the isomorphism is level-preserving.*

Hence if $(Q, *)$ is a graded structure isomorphic to A, the isomorphism is one-one onto between Q_n and A_n; thus Q_n must have a_n elements. There is a partial converse to this:

2.6 THEOREM. *Let $(Q, *)$ be a graded structure in which factoring is possible. Suppose Q_n has a_n elements, for all n. Then factoring is unique and $(Q, *)$ is isomorphic to A.*

Proof. We must merely observe that for each $n>1$,

$$*: \bigcup(Q_k \times Q_{n-k}) \to Q_n \qquad (\text{union for } 1 \leqq k \leqq n-1)$$

is one-one onto. It is onto by hypothesis, since factoring is possible. However, Q_n has a_n elements and the union on the left has

$$\sum a_k a_{n-k} = a_n \qquad (\text{sum for } 1 \leqq k \leqq n-1)$$

elements by the standard recurrence formula for Catalan numbers. Hence to be onto Q_n, the given map must be one-one. Now since factors of an element of Q_n must lie in some $Q_k \times Q_{n-k}$, factoring is unique.

2.7. *Remark.* A line of argument comparable to the above is possible for some other combinatorial problems. For instance, let B_n be the collection of distinct ways of introducing parentheses in a product of n identical terms under an operation presumed commutative but nonassociative. Let b_n denote the number of elements of B_n. It is easy to check that $b_1=b_2=b_3=1$ and $b_4=2$ [since $(xx)(xx) \neq ((xx)x)x$]. An element of B may be factored uniquely "up to commutativity" and thus it may be established that

$$b_n = \sum b_k b_{n-k} + c \qquad (\text{sum for } 1 \leqq k < n/2),$$

where $c=0$ if n is odd, $c=\frac{1}{2}b_{n/2}(b_{n/2}+1)$ if n is even. Note that if M were an anti-automorphism of this structure, $M(cd)=M(d)M(c)=M(c)M(d)$, so M would be an automorphism and thus the identity map. If this fact is taken into account, it is easy to rephrase 2.1 through 2.6 for this structure. Some history of these numbers, together with a list of combinatorial problems they solve, appears in Becker [**2**].

3. Dissections of a polygon. We regard E_1 as consisting of a (trivial) 2-gon; E_2 contains the triangle, with no diagonals. By Theorem 2.4, we may provide a one-one onto map $e: E_n \to A_n$, and thus prove that E_n has a_n elements, by defining a graded multiplication in E for which factoring is unique. We give the same operation as in [**1**], somewhat differently phrased. Let B and C denote

two dissections of an $(n+1)$-gon and an $(m+1)$-gon respectively; consider each to have one edge marked as base. Take another edge N in the plane; translate B, C, and the new edge to form a triangle, whose sides in clockwise order are (N) (base of B) (base of C). The result is combinatorially a triangulated $(n+m+1)$-gon with base N. Figure 1 illustrates a typical B, C, and their product $B\cdot C$.

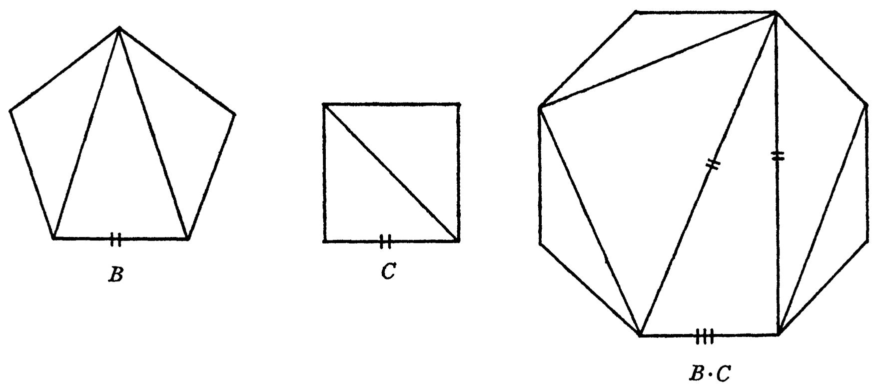


FIG. 1

We now observe that this product, denoted by $\cdot$ (usually omitted), has unique factorization. Given an element of E_n, $n>1$, with a base specified, the base belongs to a unique triangle. Erase the base and let the other two sides of its triangle be the bases for their respective subpolygons. Thus Fig. 1. B is the product of two triangles; Fig. 1. C is the product of a trivial 2-gon by a triangle; any polygon may be factored uniquely except the trivial 2-gon of E_1.

By Theorem 2.4, there is now a unique isomorphism $e:E\to A$. We describe it as follows: beginning at the center of the base, proceed clockwise around the perimeter. Write down a parenthesis "(" whenever you pass an end of some diagonal whose other end you have not yet passed. Write down an "a" whenever you pass a midpoint of any side except the base. Write down a parenthesis ")" whenever you pass an end of a diagonal whose other end you have passed previously. Thus Fig. 1. $B\cdot C$ maps to $((aa)\ (aa))\ (a(aa))$ while B and C map to $(aa)\ (aa)$ and $a(aa)$ respectively.

Let $\hat{M}$ denote reflection in the perpendicular bisector of the base, and ρ_n denote counterclockwise rotation through $2\pi/(n+1)$ radians. Now $\rho:E\to E$ may be defined by $\rho(B)=\rho_n(B)$ for B in E_n. Let $I:E\to E$ denote the identity map. The following lemmas are immediate.

3.1. LEMMA. *The relations* $\hat{M}^2=I$, $\rho_n^{n+1}=I$, *and* $\hat{M}\rho=\rho^{-1}\hat{M}$ *hold.*

3.2. LEMMA. *The order of* ρ_n *is* 1 *for* $n=1$ *and* 2, *is* 2 *for* $n=3$, *and is* $n+1$ *for* $n>3$. *Hence* ρ *has infinite order.*

Proof. The dissection D of the $(n+1)$-gon in which all $n-2$ diagonals begin at the left end of the base [hence, $e^{-1}((((aa)a)\ \cdots\ )a)$] has for $n>3$ the property

that D, $\rho_n D$, $\rho_n^2 D$, $\cdots$, $\rho_n^n D$ are distinct. Hence the order of $\rho_n \geqq n+1$ for $n>3$. The upper bound is given in 3.1, and the cases $n=1$, 2, 3 follow by inspection.

We are now able to observe that ρ_n, $n>3$, cannot preserve *any* reasonable graded multiplication on E.

3.3. THEOREM. *Let $*: E\times E\to E$ be any graded multiplication for which factoring is possible. Let $\tau: E\to E$ be any level-preserving map such that $\tau|E_k=\rho_k$ for at least one $k\geqq 4$. Then neither $\tau(B*C)=\tau(B)*\tau(C)$ nor $\tau(B*C)=\tau(C)*\tau(B)$ can hold for all B and C in E.*

Proof. By Theorem 2.6, $(E, *)$ is isomorphic to A. Thus τ induces a level-preserving map τ_A of A whose order is not less than $k+1$. Hence τ_A is not I or M, and cannot preserve multiplication.

Let $(E, \cdot)$ be the previously defined binary structure on E, and $e: (E, \cdot)\to A$ the unique isomorphism. It is clear by inspection that $\hat{M}(B\cdot C)=\hat{M}(C)\cdot\hat{M}(B)$; hence $\hat{M}$ is the unique anti-automorphism of $(E, \cdot)$. Now $e\hat{M}e^{-1}$ is an anti-automorphism of A, so $e\hat{M}e^{-1}=M$. Let ρ_A denote $e\rho e^{-1}$.

3.4. PROPOSITION: *The relations $M^2=I$ and $M\rho_A=\rho_A^{-1}M$ hold. Also $\rho_A|A_n$ has order $n+1$ for $n\geqq 4$ and $\rho_A: A\to A$ has infinite order.*

Proof. Immediate; for $M\rho_A=e\hat{M}e^{-1}e\rho e^{-1}=e\hat{M}\rho e^{-1}=e\rho^{-1}\hat{M}e^{-1}=\rho_A^{-1}\hat{M}$, and similarly for the other statements.

Since we cannot hope to find maps of A which preserve multiplication, the collection of maps which anti-commute with M (i.e., maps $\tau: A_n\to A_n$ such that $M\tau=\tau^{-1}M$) may appear worthwhile to study. We tentatively call such maps **rotations.** Thus ρ_A is a rotation of infinite order on A and of order $n+1$ on A_n for $n\geqq 4$.

We illustrate with a rotation of order 2, i.e., a map $\beta: A\to A$ such that $\beta M=M\beta$. Define $\beta: A\to A$ by the rules $\beta(a)=a$ and $\beta(bc)=M(\beta(b)\beta(c))$.

3.5. PROPOSITION. *We have $\beta^2=M^2=(M\beta)^2=I$, so $\{I, M, \beta, M\beta\}$ is the noncyclic four-group.*

Proof. Clearly $\beta(a)=M(a)=(M\beta)^2(a)=a$, so the relations desired hold on A_1. Suppose they hold on A_k, for $k\leqq n$, and let bc be in A_{n+1}. Now

$$\begin{aligned}\beta^2(bc) &= \beta M(\beta(b)\beta(c)) = \beta(M\beta(c)M\beta(b))\\ &= M(\beta M\beta(c)\beta M\beta(b)) = M\beta M\beta(b)M\beta M\beta(c)\end{aligned}$$

which is bc by the induction hypothesis that $M\beta M\beta=I$. Similarly

$$\begin{aligned}M\beta M\beta(bc) &= M\beta M[M(\beta(b)\beta(c))] = M\beta(\beta(b)\beta(c))\\ &= MM(\beta^2(b)\beta^2(c)) = bc,\end{aligned}$$

by the hypothesis $\beta^2=I$.

We observe that I, M, β, $M\beta$ are distinct since $((aa)a)a$ is carried by them to $((aa)a)a$, $a(a(aa))$, $a((aa)a)$, and $(a(aa))a$ respectively.

4. Sequences of plus and minus. Given that S_n has the same number of

elements as A_n, which we shall prove shortly, we have the following surprising result:

4.1. THEOREM. *Let $*$ be any graded multiplication on S for which factoring is possible. Then $\lambda: S \to S$ does not preserve multiplication.*

Proof. By Theorem 2.6, $(S, *)$ is isomorphic to A. By examples in 1.4 and 1.3 respectively, λ is carried to neither M nor I. Hence λ is a map from $(S, *)$ onto $(S, *)$ which is neither an automorphism nor an anti-automorphism.

We now introduce a graded multiplication is S, making it isomorphic to A and thus establishing that S_n has the same number of elements as A_n. Note that S_1 consists of the empty sequence $(\ )$ and S_2 of the single sequence $(+1, -1)$. If s_n is in S_n and s_m is in S_m, define $s_n * s_m$ to be $(+1, s_n, -1, s_m)$. For instance

$$(+1, -1, +1, -1) * (+1, -1) = (+1, +1, -1, +1, -1, -1, +1, -1).$$

If s_n has $2n-2$ terms and s_m has $2m-2$ terms, then $s_n * s_m$ has $2(n+m)-2$ terms as desired.

We now establish that factoring is unique in $(S, *)$. Suppose $(x_1, x_2, \cdots, x_{2n-2})$ is in S_n. Let $2k$ be the unique integer between 2 and $2n-2$ such that $x_1 + x_2 + \cdots + x_{2k} = 0$ and $x_1 + x_2 + \cdots + x_j > 0$ for $j < 2k$. Then $x_{2k} = -1$ and

$$(x_1, x_2, \cdots, x_{2n-2}) = (x_2, \cdots, x_{2k-1}) * (x_{2k+1}, \cdots, x_{2n-2}),$$

where the first factor is in S_k and the second in S_{n-k}. This way of factoring is the only one, since if

$$(+1, x_2, \cdots, x_{2k-1}, -1, x_{2k+1}, \cdots, x_{2n-2}) = (x_2, \cdots, x_{2k-1}) * (x_{2k+1}, \cdots, x_{2n-2}),$$

then we must have $x_1 + x_2 + \cdots + x_{2k} = +1 + (x_2 + \cdots + x_{2k-1}) - 1 = +1 + 0 - 1 = 0$ and $x_1 + x_2 + \cdots + x_j = +1 + (x_2 + \cdots + x_j) \geqq +1 + 0 > 0$ for $j < 2k$.

This definition of multiplication seems noticeably skewed. It is reasonable to introduce another one:

$$s_n \circ s_m = (s_n, +1, s_m, -1).$$

The same considerations as for $*$ apply, and we have the relations

$$\begin{aligned} \lambda(s_n \circ s_m) &= \lambda(s_n, +1, s_m, -1) = (+1, \lambda s_m, -1, \lambda s_n) \\ &= \lambda s_m * \lambda s_n, \\ \lambda(s_n * s_m) &= \cdots = \lambda s_m \circ \lambda s_n. \end{aligned}$$

Let $f_*: (S, *) \to A$ and $f_\circ: (S, \circ) \to A$ denote the unique isomorphism from $(S, *)$ and $(S, \circ)$ respectively to A. Let σ denote $f_\circ^{-1} f_*: (S, *) \to (S, \circ)$.

4.2. THEOREM. *The mapping σ is the unique isomorphism from $(S, *)$ to $(S, \circ)$; λ is the unique anti-isomorphism from $(S, \circ)$ to $(S, *)$. Hence $\lambda\sigma$ and $\sigma\lambda$ are the unique anti-automorphisms of $(S, *)$ and $(S, \circ)$ respectively.*

Proof. Immediate from the uniqueness theorems of Section 2 and the formula $\lambda(s_n \circ s_m) = \lambda s_m * \lambda s_n$.

Since $\lambda\sigma$ is an anti-automorphism, $(\lambda\sigma)^2 = 1$. Hence $\lambda\sigma = \sigma^{-1}\lambda$ and $\lambda\sigma^{-1} = \sigma\lambda$. It also follows that σ is a rotation of $(S, *)$, as we have defined that term; for the anti-automorphism of $(S, *)$ is $\lambda\sigma$ and $(\lambda\sigma)\sigma = \sigma^{-1}(\lambda\sigma)$ as required.

By contrast, $(\lambda\sigma)\lambda = \sigma^{-1} \neq \sigma = \lambda(\lambda\sigma)$, so λ is not a rotation by our definition. This may, however, indicate a failure of the definition, for this is *not* an intrinsic property of λ.

4.3. PROPOSITION. *There exists a binary operation $f: S \times S \to S$ such that (S, f) is isomorphic with A, and if $\hat{M}: (S, f) \to (S, f)$ is the unique anti-automorphism, then $\lambda\hat{M} = \hat{M}\lambda$.*

Proof. We must merely show that such an operation f can be chosen, out of the very large number of operations on S for which factoring is unique. We shall first describe a map $\hat{M}: S \to S$.

STEP 1. Let $\hat{M}(s) = \lambda(s)$ for all s in S for which $\lambda(s) \neq s$.

STEP 2. If n is odd, S_n has an even number of elements by recurrence (2) of Section 1. Organize those not used in Step 1 into pairs (s, t), and let $\hat{M}(s) = t$ and $\hat{M}(t) = s$.

STEP 3. In S_{2n}, choose arbitrarily a_n elements s for which $\lambda(s) = s$, and for them let $\hat{M}(s) = s$. At least a_n such elements exist, since if t is in S_n, then $(t, +1, -1, \lambda t)$ is such an s. However, $a_{2n} - a_n$ is even (since $a_n^2 - a_n$ is even), so the remaining elements for which $\lambda(s) = s$ may be paired as in Step 2.

We now describe an operation $f: S \times S \to S$:

STEP 4. For s in S_n, let $f(s, \hat{M}(s))$ be an element t of S_{2n} with $\hat{M}(t) = t$. This can be done in a one-one onto fashion since there are a_n pairs $(s, \hat{M}(s))$.

STEP 5. Each other pair (s, t) in $S \times S$ has an associated pair $(\hat{M}(t), \hat{M}(s))$ distinct from itself. Map (s, t) and $(\hat{M}(t), \hat{M}(s))$ to an arbitrarily chosen pair $f(s, t)$ and $\hat{M}(f(s, t))$ of suitable grade.

It is now easy to check that $\hat{M}$ is the unique anti-automorphism of (S, f) and that $\lambda\hat{M} = \hat{M}\lambda$.

We examine briefly the map in A corresponding to σ; let $\mu = f_* f_{\circ}^{-1}: A \to A$.

4.4. PROPOSITION. *We have $M\mu M\mu = I: C \to C$; hence $\mu M = M\mu^{-1}$.*

Proof. Clearly $M = f_*\lambda\sigma f_*^{-1}$, since $f_*\lambda\sigma f_*^{-1}$ is an anti-automorphism of A. Since $\mu = f_*\sigma f_*^{-1}$, we have $M\mu M\mu = f_*\lambda^2\sigma^{-2}\sigma^2 f_*^{-1} = I$.

Hence μ is a rotation of A. By observation of its behaviour on A_6, this rotation must have order at least 30; it is almost certainly infinite. We observe here that neither μ nor ρ_A is a power of the other, since

$$\mu(((aa)a)a) = f_* f_{\circ}^{-1}(((aa)a)a) = f_*(+1, -1, +1, -1, +1, -1) = a(a(aa))$$

and $\mu(a(a(aa))) = ((aa)a)a$, but repeated application of $\rho_A | A_4$ to $((aa)a)a$ yields in turn all five elements of A_4, since the five Euler triangulations of the pentagon are all rotated images of each other (one appears in Fig. 1.*B*).

References

1. Ronald Alter and Thaddeus B. Curtz, On binary nonassociative products and their relation to a classical problem of Euler, University of Kentucky, Computer Science Dept.
2. H. W. Becker (Discussion of problem 4277), this MONTHLY, 56 (1949) 697–699.
3. W. G. Brown, Historical note on a recurrent combinatorial problem, this MONTHLY, 72 (1965) 973–977.
4. A. Cayley, On the analytical forms called trees, Philosophical Magazine, 18 (1859) 374–78.
5. H. Dörrie, 100 Great Problems of Elementary Mathematics, Dover, New York, 1965.
6. Paul Erdös and Irving Kaplansky, Sequences of plus and minus, Scripta Math., 12 (1946) 73–75.
7. Haya Friedman and Dov Tamari, Problèmes d'associativité: une structure de treillis finis induite par une loi demi-associative, J. Combinatorial Theory, 2 (1967) 215–242.
8. Lajos Takács, Applications of ballot theorems in the theory of queues, Proceedings of the Symposium in Congestion Theory, Chapter 12 (W. L. Smith and W. E. Wilkinson, eds.), University of North Carolina Press, Chapel Hill, N. C., 1965.
9. W. T. Tutte, A census of slicings, Canad. J. Math., 14 (1962) 708–722.
10. A. M. Yaglom and I. M. Yaglom, Challenging Mathematical Problems with Elementary Solutions, Vol. 1, Holden-Day, San Francisco 1964.

ARZELÀ'S DOMINATED CONVERGENCE THEOREM FOR THE RIEMANN INTEGRAL

W. A. J. LUXEMBURG, California Institute of Technology

1. Introduction. Riemann's definition ([14], p. 239) of a definite integral gave rise to a number of important developments in analysis. In the course of these developments a remarkable result due to C. Arzelà ([1], 1885) marked the beginning of a deeper understanding of the continuity properties of the Riemann integral as a function of its integrand. The result of Arzelà we have in mind is the so-called ARZELÀ DOMINATED CONVERGENCE THEOREM for the Riemann integral concerning the passage of the limit under the integral sign. It reads as follows.

THEOREM A (C. Arzelà, 1885). *Let $\{f_n\}$ be a sequence of Riemann-integrable functions defined on a bounded and closed interval $[a, b]$, which converges on $[a, b]$ to a Riemann-integrable function f. If there exists a constant $M>0$ satisfying $|f_n(x)| \leqq M$ for all $x \in [a, b]$ and for all n, then $\lim_{n\to\infty} \int_a^b |f_n(x)-f(x)|\,dx=0$. In particular,*

$$\lim_{n\to\infty} \int_a^b f_n(x)dx = \int_a^b \lim_{n\to\infty} f_n(x)dx = \int_a^b f(x)dx.$$

Usually, Arzelà's theorem is formulated as a result about term-by-term

Professor Luxemburg received his Ph.D. under A. C. Zaanen at the Technological Univ. of Delft. He spent a year on a Canadian National Research Council Fellowship, two years at the Univ. of Toronto, and joined Cal. Tech. in 1958, where he now serves as the Math. Dept. Chairman. He spent a half year leave at the Univ. of Leiden in 1965. His main research interests are classical, functional, and non-standard analysis. He edited *Applications of Model Theory* (Holt, Rinehart and Winston, 1969). *Editor.*

integration for infinite series of integrable functions. In that case, the sequence of the partial sums of the infinite series plays the role of the sequence $\{f_n\}$ in Theorem A.

Due to the development of the theory of Lebesgue integration we recognize nowadays that Arzelà's theorem rests on the countable additivity property of the Lebesgue measure. As a matter of fact, Arzelà based the proof of his theorem on the following result about systems of intervals.

THEOREM B (C. Arzelà, 1885). *Assume for each n that D_n denotes a subset of $[a, b]$ that is the union of a finite (or countably infinite) number of mutually disjoint intervals. If for each n, the sum $\ell(D_n)$ of the lengths of the intervals in each D_n satisfies $\ell(D_n) > \delta$, where $\delta > 0$, then there exists at least one point $c \in [a, b]$ which satisfies $c \in D_n$ for infinitely many n.*

If one knows that the interval function measuring the length of an interval is countably additive, then Arzelà's Theorem B follows immediately by observing that

$$\ell(\limsup_{n\to\infty} D_n) \geqq \limsup_{n\to\infty} \ell(D_n) \geqq \delta > 0.$$

In the axiomatic approach to the theory of integration, Arzelà's theorem in one form or another is taken as one of the basic axioms. For instance, in abstract measure theory (see [**18**], Chap. 2) it appears in the form of the axiom that measures are countably additive. One of the basic axioms introduced by Daniell (see [**7**], [**3**] and [**18**], Chap. 3), in his functional approach to the theory of integration, is Arzelà's theorem for decreasing sequences of functions that decrease everywhere to zero. The extension procedures of the theory of integration for abstract measures as well as for abstract integrals are such that all the axioms are preserved unter the extension. Consequently, the general dominated convergence theorem for the extended integral in the abstract theory of integration is relatively easy to prove. The situation, however, is quite different in the more elementary theories of integration. For instance, in the theory of the Riemann integral the concept of a Riemann-integrable function and the value of its integral is usually defined at the outset before the basic properties of the interval function, which measures the length of an interval, such as countable additivity or even additivity have been established. This is the main reason why Arzelà's theorem for the Riemann integral has the reputation of being difficult to prove without using results from the theory of Lebesgue measure. On the other hand, in the theory of Lebesgue integration the countable additivity property of the Lebesgue measure is one of the first basic results which are established. The Arzelà-Lebesgue dominated convergence theorem follows then rather easily. This state of affairs may account for the fact that the search for an "elementary proof", roughly meaning, independent of the theory of Lebesgue measure, for Arzelà's theorem is still on. A number of elementary proofs were published by F. Riesz [**15**] in 1917, by L. Bieberbach [**4**] and E. Landau [**10**] in 1918, by F. Hausdorff [**9**] and H. S. Carslaw [**6**] in 1927, by H. A. Lauwerier [**11**] in 1949,

by J. D. Weston [16] in 1951, by W. F. Eberlein [8] in 1957, and by the present author [12] in 1961, respectively. Incidentally, in 1897, independently of C. Arzelà, W. F. Osgood [13] rediscovered Arzelà's theorem for continuous functions.

A few words concerning the known elementary proofs seem to be in order. L. Bieberbach gave a new and more elementary proof of Theorem B, and showed once more how to derive Arzelà's theorem from Theorem B. It seems that Arzelà's original proof of Theorem B (see [1], pp. 532–537) contained a gap which he filled later (see [1], pp. 596–599). A more detailed account of Arzelà's investigations can be found in [2]. E. Landau [10] gave an elementary and short proof that Theorem B implies Theorem A, thereby improving in part, Bieberbach's proof for Arzelà's theorem. F. Riesz [15] was the first to give a real elementary proof of Arzelà's theorem for continuous functions. He based his proof on Dini's uniform convergence theorem for monotone sequences of continuous functions rather than on Theorem B. F. Hausdorff [9] showed that Dini's theorem could also be used to obtain Arzelà's theorem for Riemann-integrable functions. But Hausdorff's proof seems to contain an error, which we shall discuss in more detail below. In [6], H. S. Carslaw presents his own version of the Bieberbach-Landau proof, which he remarks had gone unnoticed until that time in the English speaking world. In a footnote in the same article ([6], p. 438) Carslaw asks whether there exists also an elementary proof of a generalization of Arzelà's theorem due to W. H. Young [17]. An affirmative answer to this question is presented in the final section of the present article. H. A. Lauwerier [11] uses a form of Egoroff's theorem, but where he refers to Jordan content he really means Lebesgue measure. From a pair of inequalities for upper and lower integrals combined with an argument which could be used to prove Theorem B, J. D. Weston [16] obtains still another proof of Arzelà's theorem. W. F. Eberlein [8] proves Arzelà's theorem for Radon measures, defined on the space of real continuous functions on a compact Hausdorff space. Eberlein's proof is completely different from the proofs we have discussed so far. It is geometric in nature in that it is based on the parallelogram law and the minimal distance property for convex sets in inner product spaces. It is strongly recommended for study to the interested reader. In [12], the present author proves Arzelà's theorem for the abstract Riemann integral. It rests on a simplified version of a modification, due to I. Amemiya, of a technique used by F. Riesz in [15].

Despite the availability of this variety of elementary proofs for Arzelà's theorem, the present author finds that in most textbooks on analysis, whose authors have chosen to treat the Riemann integral rather than the Lebesgue integral, Arzelà's theorem is not mentioned, or, if it is mentioned, it is rarely accompanied by a correct proof or by any proof at all. In view of this, we should like to make one more attempt to show that Arzelà's theorem for the Riemann integral can be proved in an elementary fashion. By searching through the literature the author discovered that his new proof is essentially the same as Hausdorff's proof published in 1927. But with the important exception that at

one point, where Hausdorff gives an incorrect inductive argument, the present author gives a simple direct argument which constitutes the main part of the proof of Lemma 2.2 below.

2. An elementary proof of Arzelà's theorem. Any proof of Arzelà's theorem depends in no small measure on how the Riemann integral is introduced. We shall assume in the rest of the paper, that a bounded real function on a bounded and closed interval is Riemann integrable if its lower Darboux integral is equal to its upper Darboux integral, and that the value of its integral is equal to the common value of its lower and upper integrals. Furthermore, we assume that the reader is familiar with the DINI UNIFORM CONVERGENCE THEOREM: *Each monotone sequence of continuous functions that converges pointwise to a continuous function on a bounded and closed interval is uniformly convergent.* We shall also use the following notation: $[a, b]$ denotes the bounded and closed interval $\{x : a \leqq x \leqq b\}$; $B[a, b]$ denotes the family of all bounded functions on $[a, b]$; $C[a, b]$ denotes the family of all continuous functions on $[a, b]$; $R[a, b]$ denotes the family of all Riemann-integrable functions on $[a, b]$;

$$\underline{\int_a^b} f(x)dx \quad \text{and} \quad \overline{\int_a^b} f(x)dx$$

denote the lower and upper Darboux integrals of a function $f \in B[a, b]$ respectively. By a step function we mean a finite linear combination of characteristic functions of intervals of finite length.

The proof of Arzelà's theorem will be based on the following two lemmas:

(2.1) LEMMA. *For each* $0 \leqq f \in B[a, b]$ *and for each* $\epsilon > 0$, *there exists a continuous function* $g \in C[a, b]$ *satisfying* $0 \leqq g \leqq f$ *and*

$$\underline{\int_a^b} f(x)dx \leqq \int_a^b g(x)dx + \epsilon.$$

Proof. From the definition of the lower integral it follows that for each $\epsilon > 0$ there exists a step function s on $[a, b]$ satisfying $0 \leqq s \leqq f$ and

$$\underline{\int_a^b} f(x)dx \leqq \int_a^b s(x)dx + \epsilon/2.$$

It is easy to see that s can be transformed into a trapezoidal function g, such that $0 \leqq g \leqq s$ and $\int_a^b s(x)dx \leqq \int_a^b g(x)dx + \epsilon/2$. Hence, there exists a continuous function $g \in C[a, b]$ satisfying $0 \leqq g \leqq f$ and

$$\underline{\int_a^b} f(x)dx \leqq \int_a^b g(x)dx + \epsilon;$$

and the proof is finished.

(2.2) LEMMA. *Let* $\{f_n\}$ *be a decreasing sequence of bounded functions on*

$[a, b]$. *If* $\lim_{n\to\infty} f_n(x) = 0$ *for all* $x \in [a, b]$, *then*

$$\lim_{n\to\infty} \underline{\int_a^b} f_n(x)dx = 0.$$

Proof. It follows from (2.1) that for each $\epsilon > 0$ and for each n, there exists a continuous function $g_n \in C[a, b]$ such that $0 \leqq g_n \leqq f_n$ and

$$\underline{\int_a^b} f_n(x)dx \leqq \int_a^b g_n(x)dx + \epsilon/2^n.$$

For each n, we set $h_n = \min(g_1, g_2, \cdots, g_n)$. Then $0 \leqq h_n \leqq g_n \leqq f_n$, $h_n \in C[a, b]$, and the sequence $\{h_n\}$ decreases to zero everywhere on $[a, b]$. Hence, by Dini's uniform convergence theorem, the sequence $\{h_n\}$ converges uniformly to zero on $[a, b]$, and consequently $\lim \int_a^b h_n(x)dx = 0$. The proof of the lemma will be finished if the following inequalities are established. For each n,

$$0 \leqq \underline{\int_a^b} f_n(x)dx \leqq \int_a^b h_n(x)dx + \epsilon(1 - (1/2^n)). \tag{2.3}$$

To this end, we shall first prove the following inequalities. For each n,

$$0 \leqq g_n \leqq h_n + \sum_{i=1}^{n-1} (\max(g_i, \cdots, g_n) - g_i). \tag{2.4}$$

The inequalities (2.4) follow easily by observing that for each $1 \leqq i \leqq n$,

$$\begin{aligned} 0 \leqq g_n = g_i + (g_n - g_i) &\leqq g_i + (\max(g_i, \cdots, g_n) - g_i) \\ &\leqq g_i + \sum_{i=1}^{n-1} (\max(g_i, \cdots, g_n) - g_i), \end{aligned}$$

so (2.4) follows. From $\max(g_i, \cdots, g_n) \leqq \max(f_i, \cdots, f_n) = f_i$ it follows that

$$\underline{\int_a^b} f_i(x)dx \geqq \int_a^b (\max(g_i, \cdots, g_n) - g_i)dx + \int_a^b g_i(x)dx,$$

so

$$\int_a^b (\max(g_i, \cdots, g_n) - g_i)dx \leqq \underline{\int_a^b} f_i(x)dx - \int_a^b g_i(x)dx \leqq \epsilon/2^i$$

for $i = 1, 2, \cdots, n$. Hence, by (2.4), for each n,

$$\int_a^b g_n(x)dx \leqq \int_a^b h_n(x)dx + \sum_{i=1}^{n-1} \epsilon/2^i = \int_a^b h_n(x)dx + \epsilon(1 - (1/2^{n-1})). \tag{2.5}$$

Finally, $\underline{\int_a^b} f_n(x) \leqq \int_a^b g_n(x)dx + \epsilon/2^n$ and (2.5) imply (2.3), and the proof is finished.

REMARK. The reader does well to observe that Lemma 2.2 for Riemann integrable functions is already Arzelà's theorem for monotone sequences.

We shall now turn to the proof of Arzelà's theorem.

To this end, it is no loss in generality to assume that $0 \leqq f_n(x) \leqq M$ for all n and for all $x \in [a, b]$ and that $f(x) = \lim_{n\to\infty} f_n(x) = 0$ for all $x \in [a, b]$. For each n, and for each $x \in [a, b]$, we set $p_n(x) = \sup_{k \geqq 0} (f_{n+k}(x))$. Then $0 \leqq f_n \leqq p_n$ and the sequence $\{p_n\}$ decreases everywhere to zero on $[a, b]$. Indeed, $0 = \lim_{n\to\infty} f_n(x) = \limsup_{n\to\infty} f_n(x) = \lim_{x\to\infty} p_n(x)$ for all $x \in [a, b]$. Hence, by Lemma 2.2,

$$\lim_{n\to\infty} \underline{\int_a^b} p_n(x)dx = 0,$$

and so, $0 \leqq \lim_{n\to\infty} \underline{\int_a^b} f_n(x)dx \leqq \lim_{n\to\infty} \underline{\int_a^b} p_n(x)dx = 0$, that is, $\lim_{n\to\infty} \underline{\int_a^b} f_n(x)dx = 0$, and the proof is finished.

One may perhaps feel that the above proof follows too closely the corresponding proof of the dominated convergence theorem for the Lebesgue integral and that an elementary proof of Arzelà's theorem should deal with Riemann integrable functions only. We will show that this is possible. It is clear that Lemma 2.1 can be shown to hold for Riemann integrable functions in exactly the same way. With respect to Lemma 2.2 the situation is somewhat different. For Riemann integrable functions we can avoid the inequalities (2.4) and prove the inequalities (2.3) directly as follows: We shall use the same notation as in the proof of Lemma 2.2 with the extra hypothesis that the functions f_n are Riemann integrable. Since the sequence $\{f_n\}$ is decreasing, we have for each n, $f_n = \min(f_1, \cdots, f_n)$, and so

$$0 \leqq f_n - h_n = \min(f_1, \cdots, f_n) - \min(g_1, \cdots, g_n) \leqq \sum_{i=1}^{n} (f_i - g_i).$$

Hence, for each n,

$$0 \leqq \int_a^b f_n(x)dx - \int_a^b h_n(x)dx \leqq \sum_{i=1}^{n} \int_a^b (f_i(x) - g_i(x))dx \leqq \sum_{i=1}^{n} \epsilon/2^i$$
$$= \epsilon(1 - (1/2^n)).$$

Since the lower integral is not subadditive but superadditive, the above method cannot be used to prove Lemma 2.2. That is why we had to introduce the inequalities (2.4) to obtain (2.3). It is also this point, where Hausdorff's proof is in error.

Having established Lemma 2.2 for Riemann integrable functions, we have, in fact, proven Arzelà's theorem for monotone sequences. The next question which we have to answer is whether we can deduce Arzelà's theorem directly from its special case for monotone sequences. It is not without interest that this is indeed true. This fact is contained in the paper by F. Riesz [15] as well as in the paper by W. F. Eberlein [8]. For the sake of completeness we shall show how this can be done. In order to bring out more dramatically that Arzelà's theorem is a logical consequence from the special case for monotone sequences, we shall adopt the following abstract setting:

Let X be a non-empty set, and let L be a linear space of real functions defined on X, satisfying $f \in L$ implies $|f| \in L$. The latter condition implies that for every

finite set of elements $\{f_1, \cdots, f_n\}$ of L, $\max(f_1, \cdots, f_n) \in L$ and min $(f_1, \cdots, f_n) \in L$. A positive linear functional I on L is called an **integral** whenever I has the following property:

(2.6) *If* $0 \leqq f_n \in L$ *for each* n, *and the sequence* $\{f_n\}$ *decreases to zero everywhere on* X, *then* $\lim_{n\to\infty} I(f_n) = 0$.

It is obvious that (2.6) is Lemma 2.2 for I and L. We shall now show that (2.5) implies the following (abstract) Arzelà-type theorem:

(2.7) THEOREM. *Let* $f \in L$ *be the limit of an everywhere on* X *convergent sequence* $\{f_n\}$ *of elements of* L. *If there exists an element* $0 \leqq g \in L$ *satisfying* $|f_n(x)| \leqq g(x)$ *for all* $x \in X$ *and for all* n, *then for every integral* I *on* L *we have* $\lim_{n\to\infty} I(|f_n - f|) = 0$. *In particular*, $\lim_{n\to\infty} I(f_n) = I(f)$.

Proof. By considering the sequence $\{|f_n(x) - f(x)|\}$, which satisfies $|f_n(x) - f(x)| \leqq |f(x)| + g(x)$ for all $x \in X$ and for all n, where $|f| + g \in L$, we may assume without loss of generality that $f_n(x) \geqq 0$ for all $x \in X$ and for all n and that $f(x) = 0$ for all $x \in X$. Following F. Riesz [15], we set $g_{n,m} = \max(f_n, f_{n+1}, \cdots, f_m)$ for each pair of indices $m \geqq n$. Then $0 \leqq g_{m,n} \in L$ and $0 \leqq g_{m,n} \leqq g$ for all $m \geqq n$. Furthermore, for each n, the sequence $\{g_{m,n}\}$, and consequently, the sequence $\{I(g_{m,n})\}$ is increasing and bounded in $m > n$. Hence, for each $\epsilon > 0$ and for each n there exists an index $m_n > n$ such that $m_n < m_{n+1}$ and

$$0 \leqq I(g_{n,k}) - I(g_{n,m_n}) \leqq \epsilon/2^n, \tag{2.8}$$

for all $k \geqq m_n$. For the sake of simplicity we set $u_n = g_{n,m_n}$. Then

$$0 \leqq \limsup_{n\to\infty} u_n(x) \leqq \limsup_{n\to\infty} f_n(x) = \lim_{n\to\infty} f_n(x) = 0$$

for all $x \in X$ implies that $\lim_{n\to\infty} u_n(x) = 0$ for all $x \in X$. If we apply the inequalities (2.4) to the sequence $\{u_n\}$, we obtain for each n,

$$0 \leqq f_n \leqq u_n \leqq \min(u_1, \cdots, u_n) + \sum_{i=1}^{n-1} (\max(u_i, \cdots, u_n) - u_i). \tag{2.9}$$

Since $\max(u_i, \cdots, u_n) - u_i = \max(f_i, \cdots, f_{m_n}) - u_i = g_{i,m_n} - g_{i,m_i}$, and $m_n > m_i$ for $n > i$, we conclude that $I((\max(u_i, \cdots, u_n) - u_i)) < \epsilon/2^i$ for $1 \leqq i \leqq n$, and so, by (2.9), for each n,

$$0 \leqq I(f_n) \leqq I(\min(u_1, \cdots, u_n)) + \epsilon(1 - (1/2)^{n-1}). \tag{2.10}$$

From $\lim_{n\to\infty} u_n(x) = 0$ for all $x \in X$, it follows that the sequence $\{\min(u_1, \cdots, u_n)\}$ decreases everywhere to zero on X. Hence, by hypothesis, $\lim_{n\to\infty} I(\min(u_1, \cdots, u_n)) = 0$, and finally, using (2.10), we obtain that $\lim_{n\to\infty} I(f_n) = 0$, and the proof is finished.

REMARK. It is not difficult to convince oneself that the above proofs can be so modified as to obtain Arzelà's theorem for Riemann integrable functions of several variables. For the Riemann-Stieltjes integral, Arzelà's theorem also holds provided it is introduced in such a way that (2.1) and (2.2) hold. For this

purpose, it is necessary and sufficient that the Stieltjes measure for intervals is defined in such a way that it is a countably additive interval function. In that case, the proofs of (2.1) and (2.2) remain the same. Conversely, (2.2) implies the countable additivity property of the Stieltjes measure.

3. Fatou's lemma for the Riemann integral. In the theory of Lebesgue integration, Fatou's lemma plays an important role. The analogous result for the Riemann integral will follow easily from the following lemma:

(3.1) LEMMA. *Let* $0 \leqq f \in R[a, b]$ *be the limit of an everywhere convergent sequence* $\{f_n\}$ *of non-negative Riemann integrable functions on* $[a, b]$. *Then*

$$\lim_{n\to\infty} \int_a^b (f(x) - f_n(x))^+ dx = 0,$$

where $(f(x) - f_n(x))^+ = \max(f(x) - f_n(x), 0)$ *for all* $x \in [a, b]$.

Proof. Since the functions f_n, f are non-negative it follows that $f(x) - f_n(x) \leqq f(x)$ for all $x \in [a, b]$. Hence, $(f(x) - f_n(x))^+ \leqq f(x)$; and the result follows from Arzelà's theorem.

(3.2) THEOREM. (Fatou's lemma for the Riemann integral). *Under the same hypotheses of* (3.1), *we have*

$$0 \leqq \int_a^b f(x)dx \leqq \liminf_{n\to\infty} \int_a^b f_n(x)dx.$$

Proof. Observe that $f = (f - f_n) + f_n \leqq (f - f_n)^+ + f_n$ for all n. Hence, by (3.1),

$$\int_a^b f(x)dx \leqq \liminf_{n\to\infty} \left(\int_a^b (f(x) - f_n(x))^+ dx + \int_a^b f_n(x)dx \right) = \liminf_{n\to\infty} \int_a^b f_n(x)dx,$$

and the proof is finished.

From (3.1) we can also deduce the following result supplementing Fatou's lemma:

(3.3) THEOREM. *Under the same hypotheses of* (3.1), *and* $\int_a^b f(x)dx = \lim_{n\to\infty} \int_a^b f_n(x)dx$, *we have* $\lim_{n\to\infty} \int_a^b |f(x) - f_n(x)| dx = 0$.

Proof. From (3.1), the new hypothesis, and

$$f(x) - f_n(x) = (f(x) - f_n(x))^+ - (f(x) - f_n(x))^-$$

for all n, it follows that $\lim_{n\to\infty} \int_a^b (f(x) - f_n(x))^- dx = 0$, where $(f(x) - f_n(x))^- = \max(-(f(x) - f_n(x)), 0)$, $x \in [a, b]$. Hence

$$\lim_{n\to\infty} \int_a^b |f(x) - f_n(x)| dx$$
$$= \lim_{n\to\infty} \left(\int_a^b (f(x) - f_n(x))^+ dx + \int_a^b (f(x) - f_n(x))^- dx \right) = 0,$$

and the proof is finished.

4. W. H. Young's extension of Arzelà's theorem. In [17], Test 6, p. 316, W. H. Young gave an extension of Arzelà's theorem for the Lebesgue integral, which appears as a problem about term-by-term integration in [5], Example 22, p. 144. In [6], Carslaw asks whether there exists a simple proof for Young's result. Since the result of Young is interesting in itself, we shall present it here supplied with an elementary proof.

(4.1) THEOREM (W. H. Young). *Assume f_n, g_n, and $h_n \in R[a, b]$ for each n and that the sequences $\{f_n\}$, $\{g_n\}$ and $\{h_n\}$ converge everywhere on $[a, b]$ to the Riemann integrable functions f, g, and h, respectively. If $h_n \leqq f_n \leqq g_n$ for each n, and passage of the limit under the integral sign for the sequences $\{g_n\}$, $\{h_n\}$ is possible, that is, $\lim_{n\to\infty} \int_a^b g_n(x)dx = \int_a^b \lim_{n\to\infty} g_n(x)dx = \int_a^b g(x)dx$ and $\lim_{n\to\infty} \int_a^b h_n(x)dx = \int_a^b \lim_{n\to\infty} h_n(x)dx = \int_a^b h(x)dx$, then the same holds for the sequence $\{f_n\}$, that is, $\lim_{n\to\infty} \int_a^b f_n(x)dx = \int_a^b \lim_{n\to\infty} f_n(x)dx = \int_a^b f(x)dx$.*

For the case that $h_n(x) = -M$ and $g_n(x) = M$ for all n and for all $x \in [a, b]$, where M is a positive constant, Young's theorem reduces to Arzelà's theorem.

We shall now show that Young's theorem follows from Arzelà's theorem, Theorem 3.3 and the following lemma:

(4.2) LEMMA. *If $0 \leqq u$, v, and $w \in R[a, b]$ satisfy $0 \leqq u \leqq v + w$, then u can be written in the form $u = u_1 + u_2$, where $0 \leqq u_1 \leqq v$, $0 \leqq u_2 \leqq w$, and u_1, $u_2 \in R[a, b]$.*

Proof. Let $u_1 = \min(u, v)$ and let $u_2 = u - u_1$. Then $0 \leqq u_1 \leqq v$ and $0 \leqq u_2 = u - u_1 = \min(u, v+w) - \min(u, v) \leqq v + w - v = w$. Since $\min(u, v) = \frac{1}{2}(u + v - |u - v|)$ it follows that $u_1 \in R[a, b]$, and so also $u_2 = u - u_1 \in R[a, b]$, finishing the proof.

We shall now turn to the proof of Young's theorem. From $h_n \leqq f_n \leqq g_n$ it follows that $0 \leqq f_n - h_n \leqq g_n - h_n = g_n - h_n - (g - h) + (g - h)$. Observe that the sequence $\{g_n - h_n\}$ satisfies the hypotheses of Theorem 3.3, and so, by setting $u_n = |g_n - h_n - g + h|$ for each n, we conclude that

$$\lim_{n\to\infty} \int_a^b u_n(x)dx = 0. \tag{4.3}$$

Then $0 \leqq f_n - f - h_n + f \leqq u_n + |g - h|$ for all n. Hence, from (4.1) it follows that $0 \leqq f_n - f - h_n + f = v_n + w_n$ for all n, where $0 \leqq v_n$, $w_n \in R[a, b]$, $0 \leqq v_n \leqq u_n$, and $0 \leqq w_n \leqq |g - h|$ for all n. Since $\lim_{n\to\infty} u_n(x) = 0$ for all $x \in [a, b]$ it follows that $\lim_{n\to\infty} w_n(x) = f(x) - h(x)$ for all $x \in [a, b]$. Then $0 \leqq w_n \leqq |g - h| \in R[a, b]$ for all n implies, by Arzelà's theorem,

$$\lim_{n\to\infty} \int_a^b w_n(x)dx = \int_a^b f(x)dx - \int_a^b h(x)dx. \tag{4.4}$$

Hence, by (4.3) and (4.4) we have

$$\lim_{n\to\infty}\left(\int_a^b (f_n(x) - f(x))dx - \int_a^b h_n(x)dx + \int_a^b f(x)dx\right) = \int_a^b f(x)dx - \int_a^b h(x)dx. \tag{4.5}$$

From the hypothesis $\lim_{n\to\infty} \int_a^b h_n(x)dx = \int_a^b h(x)dx$ and (4.5) it follows finally that $\lim_{n\to\infty} \int_a^b (f_n(x) - f(x))dx$ exists and is equal to zero, and the proof is finished.

REMARK. The reader will have no difficulty in showing that if, in addition, the sequences $\{g_n\}$, and $\{h_n\}$ satisfy $\lim_{n\to\infty} \int_a^b |h_n(x) - h(x)| dx = 0$ and

$$\lim_{n\to\infty} \int_a^b |g_n(x) - g(x)| dx = 0,$$

then the sequence $\{f_n\}$ also has the property $\lim_{n\to\infty} \int_a^b |f_n(x) - f(x)| dx = 0$.

Work on this paper was supported in part by NSF grant GP 23392.

References

1. C. Arzelà, Sulla integrazione per serie, Atti Acc. Lincei Rend., Rome, (4) 1 (1885), 532–537, 596–599.

2. ———, Sulle serie di funzioni, Mem. Inst. Bologna (5), 8 (1900) 131–186 and 701–744.

3. S. Banach, The Lebesgue integral in abstract spaces, note in S. Saks, Theory of the Integral, Warsaw 1933 and New York 1937, 320–330.

4. L. Bieberbach, Über einen Osgoodschen Satz aus der Integralrechnung, Math. Z., 2 (1918) 155–157.

5. T. J. I'A. Bromwich, An Introduction to the Theory of Infinite Series, sec. ed. rev., London 1926.

6. H. S. Carslaw, Term-by-term integration of infinite series, Math. Gaz., 13 (1927) 437–441.

7. P. J. Daniell, A general form of integral, Ann. of Math., 19 (1917) 279–294.

8. W. F. Eberlein, Notes on Integration I: The underlying convergence theorem, Comm. Pure Appl. Math., 10 (1957) 357–360.

9. F. Hausdorff, Beweis eines Satzes von Arzelà, Math. Z., 26 (1927) 135–137.

10. E. Landau, Ein Satz über Riemannsche Integrale, Math. Z., 2 (1918) 350–351.

11. H. A. Lauwerier, An elementary proof of the Arzelà-Osgood-Lebesgue theorem, Simon Stevin, 26 (1949) 177–179 (Dutch).

12. W. A. J. Luxemburg, The abstract Riemann integral and a theorem of G. Fichtenholz on equality of repeated Riemann integrals. IA and IB, Proc. Ned. Akad. Wetensch. Ser. A 64 (1961) 516–545 = Indag. Math., 23 (1961).

13. W. F. Osgood, Non-uniform convergence and the integration of series term-by-term, Amer. J. Math., 19 (1897) 155–190.

14. B. Riemann, Über die Darstellbarkeit einer Funktion durch eine trigonometrische Reihe, Habilitationsschrift 1854, Abhandl. Gött. Ges. Wiss., 13 (1868) Gesammelte Werke, 1892, 227–264, New York 1953.

15. F. Riesz, Über Integrationen unendlicher Folgen, Jber. D. Math. Verein., 26 (1917) 274–278.

16. J. D. Weston, Inequalities for Riemann-Stieltjes integrals, Math. Z., 54 (1951) 272–274.

17. W. H. Young, On semi-integrals and oscillating successions of functions, Proc. London Math. Soc., (2) 9 (1910) 286–324.

18. A. C. Zaanen, An Introduction to the Theory of Integration, sec. ed., Amsterdam 1967.

THE CRISES OF THE MATHEMATICAL SCIENCES, AND WHY NO ONE GROUP CAN SOLVE THEM[1]

GAIL S. YOUNG, University of Rochester

Now that I am no longer president of the Mathematical Association and cannot be impeached, I can say that the greatest mistake in American mathematics was the founding of the MAA. This took place some 55 years ago, as a result of a dispute between those members of the AMS who thought that the Society should confine its attention only to research, and a minority group who thought that the Society should also involve itself with the undergraduate teaching of mathematics. If the Society at that time had been changed to take responsibility for the development of all the various aspects of mathematics, we could have avoided the proliferation of specialized organizations in the mathematical sciences by having a proper place for these special interests in the Society. But that is not the route we took, and I now find myself writing about the desirability of cooperation between organizations, a problem that should never have arisen.

It seems to me my best approach should be in the form of an existence proof, the demonstration that there exist important problems that no one organization can cope with. I am not trying to be all-inclusive; the examples I shall discuss are ones that I myself know about and regard as important. There are readers who know of other problems that may be more important, and there may very well be readers who will say that some of my examples are not worth considering. All I can say is, these problems I happen to know about and I personally think should be solved.

1. Let me begin with one of the touchier subjects, a national information system for the mathematical sciences. That there are problems with publication we all know: exponential growth of papers, rapidly rising costs, slowly rising library expenditures, too many conferences—especially to which one is not invited. I want to write, however, not about the problems, but about the difficulties of solution.

I will say at once that I don't know whether we need a national information *system*. So far as my own needs are concerned, existing mechanisms work pretty well. Major improvement for me would consist in supplying me with a tutor, not a computer. I believe that for many of the users of mathematics our present mechanisms do not work well, but for all I know, relatively minor changes in our present mechanisms might be sufficient. For example, if Mathematical Reviews really had a comprehensive coverage, how much would this help the users? Thus I am not advocating or deploring any point of view.

However, the example illuminates several different aspects of the needs and problems of cooperation. To condense and oversimplify, I shall remind you that the American Mathematical Society, after committee study, made proposals to

[1] Based on a talk given to the Division of Mathematical Sciences, National Research Council, March 22, 1971.

OSIS, the Office of Science Information Study in NSF, for grants for rather minor but highly useful improvements in the existing system. These improvements perhaps would be of most help to the research mathematician, particularly in pure mathematics. The OSIS response was that these particular projects could be funded only as part of a larger operation—the development of a complete information system—and that other organizations must be involved.

Since a multi-organizational committee was established, it would be possible to say that this first example is one where there has been cooperation between our societies. I have not regarded the cooperation as quite satisfactory. We needed first of all a truly cooperative study of the need for such a system by a group given the trust of the mathematical organizations, and with its members free from the necessity of representing the interests of any particular organization. We were very far from that. In each individual organization, committees function on an implicit basis of trust, resting on long experience of the organization with the members of the committee, and on faith in the officers who select the committees. Without that trust, questions such as the desirability of an information system become, at the very least, much harder to analyze, as our progress shows.

Beyond that, there is another question. Did we *have* to take the decisions of OSIS that anything but the most trifling of projects must be put into an all-embracing framework? As president, I found myself taking rather simple things that MAA would like to have funding for and contorting them so that they could be justified as part of a system. If a system is needed, we—the mathematical community—should be able to decide this by ourselves. If we believe it is not needed, we should not be cut off from funding for our projects. No one of our organizations could hope to change the attitude of OSIS. (Other things are apparently causing a very welcome improvement.) Cooperative action between the societies is needed to get such policies changed, and we have no way of getting such action.

2. Here is another problem that involves all of us, though falling, strictly speaking, in the purview of only one of our organizations, the National Council of Teachers of Mathematics. As a result of work of groups like SMSG, there has been a great improvement in elementary and secondary school mathematics. I had thought that that improvement was permanent. Lately I have heard almost desperate pleas from members of NCTM for help against what seems to be an anti-intellectual assault on school mathematics. There are now a number of private corporations that are making deals with school boards to take over certain parts of the teaching of mathematics on a basis of guaranteeing an improvement in scores on standard objective tests of mathematics skills. I don't want to discuss this in any detail,[2] but I shall simply assume that what I have been told is correct, that the teaching in these programs by and large takes out all the intellectual content and replaces it by mechanical drill, and that there really

[2] Cf. CBMS Newsletter, Vol. 6, May 1971, page 1.

is a strong anti-mathematical bias back of it all. Suppose that is so; what can be done about it? In the first place, NCTM alone cannot effectively protest. They can be charged with working in their own selfish interests. Nor can, for example, AMS or MAA act alone. They can be told they don't know the problem. If this is to be stopped, it can only be done by the united protest of the mathematical community. Every one of our organizations will be damaged by deterioration of the mathematics programs in the schools.

3. A place where there has been rather effective cooperation has been the Survey Committee of CBMS, of which I am chairman. One reason that it has been effective was pure accident: the small group of us that drew up the original proposal to the Ford Foundation which carried us through five years of work, thought that we needed panels of specialists to make sure we were asking the right questions; thus, for example, we had a panel on statistics chaired by George Nicholson, Chairman of Statistics at the University of North Carolina, and with a number of other well-known statisticians among its members. I remember vividly the atmosphere of suspicion at CBMS when we first proposed getting under way. If we had not devised this system of panels, we might not have gotten approval. As it was, whatever person expressed suspicions, we could point to the existence of an appropriate panel formed of people that he could not possibly distrust. The panel mechanism proved not to be so efficient and so necessary as we had thought, but if it had been intended as a political maneuver, it would have been brilliant.

The atmosphere of suspicion and distrust also forced us to make a commitment that we would do nothing in the surveys that could be regarded as expression of opinion. Without that, we would not have had approval. This had the following serious effect. The first survey we made in 1965 was a repetition of the Lindquist survey of 1960, of undergraduate mathematics in the four-year colleges. As a result of comparing the two studies and examining predictions of Ph.D. production and of undergraduate enrollment, I became convinced then that around 1970 we would begin to feel the effects of an over-production of Ph.D.'s for teaching. I said as much at the time in several speeches. But I was not able to make remarks like this in the Survey report, because of the agreement. It was not something that was *prima facie* evident. One had to interpret the data and give it meaning. It required editorializing. Whether such a discussion would have been listened to, I can't say. Certainly, I was unable to convince the group in COSRIMS concerned with manpower that their projections were unrealistic. What has happened is more severe than I anticipated, because of other factors than mere numbers. However, if there is anything that I regard as self-evident and that now everyone should agree on, it is that we need a firm basis of data about our own field, gathered by people who understand our needs and who are free to draw conclusions from the data.

The work the committee has done up to now is by no means all that is needed. For example, we really know very little about school mathematics. The

committee had hoped to study this, but learned that the effort required was completely beyond its financial position. We know almost nothing about industrial use of mathematicians, and have no real data as to what the potentialities are, or as to what the proper sort of training should be. There again, that is a topic that turned out to be much too large for our financing. We have a grant from NSF to carry on a repetition of our undergraduate survey, but all our work exists only on a temporary basis of grant support. I cannot even say that CBMS will stay in business so that there can be a CBMS Survey Committee.

4. Decisions are made all the time in Congress and in the granting agencies concerning the mathematical sciences without proper consultation with us. Two years ago, when I was still chairman of the NRC Mathematics Division's Committee on Forms and Levels of Support, Garrett Birkhoff, as chairman of CBMS, arranged a meeting between Representative Daddario, himself, and me to talk over matters of governmental policy affecting mathematics. This was the only contact I have ever heard of between mathematicians and Congress. Representative Daddario was very interested in our remarks and obviously pleased that finally he had heard from some mathematicians. As for the granting agencies, we have left all the burden of representing the mathematical sciences on the shoulders of those mathematicians who are on the staffs of the various agencies. I have said before, and I would like to say it again, that our community owes a tremendous debt to the mathematicians in the agencies for the way they have worked for our aims without receiving much back from us in the way of thanks or recognition. But however ably they have done it, the task is more than they should be asked to undertake.

Now here again, no one organization can presume to speak for us all. Further, any one organization can be suspected of furthering its own interests at the expense of the others. One thing I have heard over and over again from administrators of various sorts is that mathematicians are always saying they are different. I think all of us would agree that among the other sciences, we indeed are different, with different needs and different problems. We cannot let policies or laws that have been set up with the active intervention of other sciences apply unchanged to us. We must have means of finding out what questions of policy are being studied and of bringing our opinion to bear.

I would like to make quite clear that I am not talking here about lobbying for our selfish interests. We have a responsibility to the country for the proper development of the mathematical sciences, and we are abdicating that responsibility if we permit important decisions concerning us to be made while we stand passively by.

5. At a distinctly more utilitarian level, it seems to me ridiculous that we are all in the publishing business, that we are all data processors, that we are all collection agencies, that we are all advertising agencies. Ridiculous or not, that is something I have decided that we shall probably have to live with. I see

no possibility of getting that level of cooperation in the mathematical community required to change this.

The several problems that I have outlined up to now have a certain objective quality; if the facts are as my summary has indicated, I suspect that there would be rather general agreement that something should be done. I want to discuss a problem, or group of problems, where I suspect there would be less agreement.

6. We are living in a world crisis, composed of many hard problems. I conjecture that if my readers wrote down a list of the ten most important problems facing this country, or the world, there would be an amazing consensus. Some of the problems—for example, the war in Viet Nam—are ones that we as mathematicians, or the organizations we represent, can do nothing much about. The solution of others will require a great deal of scientific work, and I am sure that mathematics and mathematical methods will be highly important. The problems will continue and intensify, and at some point it seems to me that we shall have to start working on them on a crash-program basis, much the way that in World War II the scientists all turned their attentions to the war. The problems simply will not go away. The demand for energy will continue to rise, bringing with it problems of thermal pollution, radioactivity, fuel consumption, etc. Other problems are not so inevitable, for example, the possible dangers from the supersonic transports.

One reason for the existence of such problems is the fact that we do not have enough scientists to go around, and society, so far, has been unwilling to pay for the necessary number of scientists. Let me explain what I mean. Take the case of thermal pollution of water by nuclear power plants. From the amount of controversy, it seems to be obvious that we must not really understand the effects. But how could we? Suppose that 25 years ago a biologist had speculated about the effect of raising the temperature of the Finger Lakes by 5°. He would certainly not have followed the speculation up, and would have continued to work on problems whose importance was obvious to him. I don't know where he would have received support from then if he had tried to follow it up. But it seems to me that in order to have answers now as to the effect of thermal pollution of large bodies of water, research would probably have had to start about then, that is, about the time that it first became apparent to anyone that nuclear power might become important. Would anyone 15 years ago have proposed to study the effect of introducing exhaust fumes into the upper atmosphere? Again, from the amount of controversy on the SST, it seems to me that we do not understand the basic atmospheric physics well enough to make predictions now. Something like the 15 years lead time would have been essential.

We are going to have to study many such problems. I think it should be a principle of technological development from now on, that part of every such development should be large-scale study of the effects of the development. That would take many, many more scientists than we now have. What is the implication of this for the mathematical scientists? I don't know, and that is what

makes this a problem that requires cooperation between the various organizations. As a university teacher and graduate chairman, I have responsibilities here that I do not know how to meet. I think that we should begin making large-scale studies of the implications of these environmental and other technological problems for our various sub-disciplines, and I regard it as a task that is beyond the reach of any one of our organizations.

7. Another conceivably controversial example arises from the employment crisis. As I said, it could have been predicted five years ago, and I think that a really wise person could have predicted it ten years ago at the time of the Gilliland report. The eschatological items I have just discussed could use up our "surplus," but let me pretend these do not exist. We are producing 1300 Ph.D.'s a year, almost all going into college and university teaching. Without major changes in our forms of teaching and in the amount of mathematics the average student takes, by around 1975 there will be 600 or less jobs open in teaching for new Ph.D.'s. We have now 65 graduate departments good enough to receive ratings in the 1969 ACE study of graduate departments[3] as opposed to 47 in 1964. There are around 175 departments giving the Ph.D. in mathematics, and more being added. I see no way of cutting down the number, or preventing new programs from starting, and indeed, I do not believe the number should be cut down. We need more trained people, and we need the research opportunities a graduate department provides. But we must have changes. At present, almost all of these departments are simply trying to do what the top five departments do, and that is a mistake.

But what *are* they to do? Again, I don't know. I can make some conjectures. For example, in 1975, we shall have something like 100,000 computers in operation. I pass over the question of how we are to staff them with competent people, but I am sure anyone in the Association for Computing Machinery will regard that as a serious problem. That is not a problem at the doctoral level. However, a large number of these computers will be in companies where their presence makes it possible to do a great deal of mathematical work on problems of the company that would be impossible without them. I am not referring to the computer as a tool for numerically solving classical engineering problems, but to the computer for solving problems of, let me say, operations analysis. Is it too much to estimate that one out of ten computers will be in such a place? If that is right, it means the existence of 10,000 jobs at about the doctoral level, more than the total number of Ph.D.'s in all the mathematical sciences, and enough to justify any number of graduate departments—if the departments take this problem seriously. I have talked to former IBM field men about this idea, and gotten an enthusiastic response, so that I believe the need probably exists.

However, all such solutions to the problem of "surplus" will require tremendous changes in attitudes in graduate departments. Perhaps our strongest departments should not change, but Southeastern X State must, if it wants a

[3] Cf. CBMS Newsletter, Vol. 6, January 1971, pages 8–9.

graduate program. Graduate work in mathematics is the province of the Society but the question of changing graduate programs in this sort of way again is beyond the wisdom of any one organization.

I think that this is enough of a sample. I should like very much to collect a list of such problems, and I am sure that many readers have thought of others that I have not been aware of. I should appreciate receiving additions to the list.

What machinery do we now have working on such problems? There are only two organizations that have any claim to speak for the mathematical sciences. One of these is the Division, and the other is the Conference Board of Mathematical Sciences. It does not seem to me that the Division is the appropriate body. The Division has a quasi-governmental status, and is, in any case, part of a larger body, the National Research Council, that determines the basic policy. This leaves CBMS. In January 1971, I completed five years of membership on CBMS, three as a member at large and two as representative of the Association. They were five years of disenchantment. I have already written about the suspicion shown at the beginning of such a non-controversial topic as the work of the Survey Committee. That, in my view, has been rather typical. I shall not give other examples, but all through my terms most activities that the Conference Board attempted to initiate were stopped by the action of one or more organizations acting either out of mistrust, or of a desire to protect the organization's selfish interests. The Board was not allowed to grow to where it could get outside support enough to have the staff required to do much. Until recently, it has been utterly dependent on the dues of the member organizations. From that standpoint it has not been cheap. The MAA has paid about 50¢ per member to CBMS, and has continued to do so while we have twice raised dues. Essentially, what organization dues have supplied has been office rent, the salary of the executive secretary, and the salary of his secretary. Anything else that has required support has been dependent on outside financing of a temporary nature.

In the past several years, various of the Board organizations, including MAA, set up committees to study their relationship with CBMS, essentially to decide whether or not to stay in. I was chairman of the MAA committee. By its nature, the MAA tends to have a larger view of problems in the mathematical sciences than most of the other organizations. I believe one can say that our committee was convinced of the theoretical desirability of some form of umbrella organization. A question that we had to face, however, was, given the existing societies—was there *any* form of organization that would be acceptable? If the answer was yes, should we give up on CBMS and attempt to establish a new umbrella organization, perhaps with a smaller membership? That was one course of action proposed to us. Would it be possible to reorganize CBMS to become really effective?

Our committee recommended to the Board of Governors at the 1970 Summer Meeting that they adopt a rather long motion embodying our conclusions. I shall not attempt to repeat it here. It had four main parts: An endorsement of the continued existence of CBMS; specific criticisms of the current form of

organization; detailed suggestions for improvement of the structure; and a directive for MAA to sponsor a meeting of the Conference Board organizations, to discuss the purposes of CBMS and to take measures leading to its improvement. The proposed meeting was held in November, 1970, at Airlie House, and I believe that it was a success. A quick summary would be to say, essentially, the MAA resolution for changes was endorsed.

One of the main reasons for our committee's decision to endorse CBMS was the greatly increased activity of the Board under the chairmanship of Garrett Birkhoff. If it had not been for his work, and for the demonstration he gave that given vigorous enough leadership, CBMS could be effective, even in its present form, MAA would probably have given up.

At the CBMS January meeting, these changes were approved, and a committee is at work studying ways to implement them. A proposal formed since the Airlie House meeting to establish a board of trustees was also accepted. This board would be composed of distinguished people having the trust of the community, and would be responsible for monitoring the work of CBMS. I believe this idea is due to Donald Thomsen. Another idea, due to Saunders MacLane, for individual members, is receiving further study, and I hope it will be accepted.

I am still not completely optimistic about the Conference Board's life. The ecumenical meeting at the Airlie House ended with harmony. But I have been surprised at how quickly distrust and acrimony reappeared.

I think I should say what I think will happen if CBMS collapses. There will certainly be attempts to form a "more selective" organization. Any such organization can have only limited objectives. It cannot effectively tackle the problems I have listed; these need all the CBMS organizations. The omitted organizations will have to become more active, in defense of their proper interests, perhaps forming a rival organization. Which will get listened to? Neither.

CORRECTION TO "SUMMATION OF THE SERIES $1^n+2^n+\cdots+x^n$ USING ELEMENTARY CALCULUS"

L. S. LEVY, University of Wisconsin

The article appeared in this MONTHLY 77 (1970) 840–847.

On page 841, Observation I: read $n>1$, not $n<1$.

On page 846: Another proof of Corollary 3 can be found in Joseph Arkin, *A function whose values are integers*, Mathematics Magazine 38(1965) pp. 196–199.

On page 847: read $C_{20}=-1222277/2310$, not $C_{20}=-122277/2310$ (thanks to R. P. Tapscott for pointing this out).

ADDENDUM TO "A PROOF OF THE NEWTON-COTES QUADRATURE FORMULAS WITH ERROR TERM"

D. R. HAYES, University of Massachusetts, and L. RUBIN, University of Hartford

Professor L. R. Bragg has pointed out to the authors that most of the simplifications that we introduced in our article (this MONTHLY, 77 (1970) 1065–1072) are contained in *The remainder terms in numerical integration formulas* (this MONTHLY, vol. 70 (1963) 70–76) by L. R. Bragg and E. B. Leach. In fact, Bragg and Leach give a very nice inductive proof of our Theorem 3.4 which is considerably simpler than our proof. We apologize for having overlooked their work. Our Theorem 2.1 together with the Bragg-Leach ideas should provide a proof of the error terms which is very well suited to an undergraduate class.

MATHEMATICAL NOTES

EDITED BY DAVID DRASIN

Manuscripts for this Department should be sent to Robert Gilmer, Department of Mathematics, Florida State University, Tallahassee, FL 32306.

ON A FUNCTIONAL EQUATION

HIROSHI HARUKI, University of Waterloo

We consider Cauchy's functional equation

$$f(x+y)=f(x)+f(y), \tag{1}$$

where f is an entire function. (In this note, x, y, and z are complex variables, while s and t represent real variables.)

Simple computations show that (1) implies equations

$$|f(s+it)| = |f(s)+f(it)|, \tag{2}$$

$$|f(x+y)+f(x-y)| = |f(x+\bar{y})+f(x-\bar{y})|. \tag{3}$$

The following theorem was proved in [1]:

THEOREM A. *If $f(z)$ is an entire function of a complex variable z and satisfies (2) for real values of s and t, then $f(z)=a\sin\alpha z$ or $f(z)=a\sinh\alpha z$ or $f(z)=az$, where a is an arbitrary complex constant and α is an arbitrary real constant.*

The purpose of this note is to solve (3), that is, to prove the following theorem by using Theorem A.

THEOREM. *If $f(z)$ is an entire function of a complex variable z and satisfies (3) for complex values of x and y, then and only then $f(z)=a\cos\alpha z+b\sin\alpha z$ or $f(z)=a\cosh\alpha z+b\sinh\alpha z$ or $f(z)=az+b$, where a, b are arbitrary complex constants and α is an arbitrary real constant.*

Theorem A will play a useful role in our derivation. We shall also use the following uniqueness theorem (cf. [2, p. 66]):

THEOREM B. *Let $f(z)$ be a nonconstant entire function and suppose that for $n=0, 1, 2, \cdots$, the equation*

$$f^{(n)}(z) = 0$$

has no solutions. Then $f(z)=\exp(Az+B)$, where A, B are complex constants with $A\neq 0$.

Proof of the Theorem.

Case (A). Suppose $f^{(n)}(z)\neq 0$ ($n=0, 1, 2, \cdots$) for all z.

By Theorem B we have

$$f(z) = \exp(Az + B), \tag{4}$$

where A, B are complex constants with $A\neq 0$.

Substituting (4) in (3), we have $|\cosh Ay|^2=|\cosh A\bar{y}|^2$, or

$$\cosh Ay \cosh \overline{Ay} = \cosh A\bar{y} \cosh \overline{A}y. \tag{5}$$

Upon expanding both sides of (5) in power series and equating the coefficients of y^2, we see that $A^2=\overline{A}^2$. Hence A is real or purely imaginary, and so by (4) we have

$$f(z) = a \cos \alpha z + ai \sin \alpha z,$$

or

$$f(z) = a \cosh \alpha z + a \sinh \alpha z,$$

where a is an arbitrary complex constant and α is an arbitrary real constant.

Case (B). Let p be the least integer ($\geqq 0$) such that $f^{(p)}(z)$ has at least one zero point in $|z|<+\infty$.

From (3) it follows that

$$|f(x + y) + f(x - y)|^2 = |f(x + \bar{y}) + f(x - \bar{y})|^2. \tag{6}$$

We next take Laplacians $\Delta=\partial^2/\partial s^2+\partial^2/\partial t^2$ of both sides of (6) p times with respect to $x=s+it$, and obtain

$$4^p|f^{(p)}(x + y) + f^{(p)}(x - y)|^2 = 4^p|f^{(p)}(x + \bar{y}) + f^{(p)}(x - \bar{y})|^2,$$

or

$$|f^{(p)}(x + y) + f^{(p)}(x - y)| = |f^{(p)}(x + \bar{y}) + f^{(p)}(x - \bar{y})|, \tag{7}$$

since, by [3], $\Delta|f|^2=4|f'|^2$.

Let z_0 be a zero point of $f^{(p)}(z)$ and

$$F(z) = f^{(p)}(z + z_0). \tag{8}$$

By (7), (8) we then have

$$(9) \qquad |F(x+y)+F(x-y)| = |F(x+\bar{y})+F(x-\bar{y})|.$$

Upon letting $x=y=\frac{1}{2}(s+it)(s,\ t$ real) in (9) and using the fact that $F(0)=f^{(p)}(z_0)=0$, we see that

$$(10) \qquad |F(s+it)| = |F(s)+F(it)|.$$

Equation (10) and Theorem A yield that

$$(11) \qquad F(z) = a \sin \alpha z,$$

or

$$(12) \qquad F(z) = a \sinh \alpha z,$$

or

$$(13) \qquad F(z) = az,$$

where a is a complex constant and α is a real constant.

It remains to show that $F(z)=f(z+z_0)$. We may assume that $a\alpha\neq 0$, and shall show that the assumption $p>0$ in (8) leads to a contradiction.

By (8) we have

$$f^{(p-1)}(z) = -\frac{a}{\alpha}\cos \alpha(z-z_0)+C,$$

or $f^{(p-1)}(z)=(a/\alpha)\cosh\ \alpha(z-z_0)+C$, or $f^{(p-1)}(z)=\frac{1}{2}a(z-z_0)^2+C$, where C is a complex constant.

Since $a\alpha\neq 0$, $f^{(p-1)}(z)$ has at least one zero point and this then contradicts the fact that

$$f^{(p-1)}(z) \neq 0 \text{ in } |z| < +\infty.$$

Hence $p=0$, and the theorem follows from (8), (11), (12), and (13).

References

1. R. M. Robinson, A curious trigonometric identity, this MONTHLY, 64 (1957) 83–85.
2. W. K. Hayman, Meromorphic Functions, Oxford Math. Monographs 1964, pp. 66–67.
3. G. Pólya and G. Szegö, Aufgaben und Lehrsätze aus der Analysis, I, Berlin-Göttingen-Heidelberg, Springer, 1954, p. 94, problem 58.

PERMUTATIONS ARRANGED AROUND A CIRCLE

EMILE ROTH, Regionales Rechenzentrum, Stuttgart, Germany

The result to be established will be introduced by an example. Let the 9-tuple $(a, a, b, b, a, c, c, b, c)$ be written clockwise around a circle. Then the ordered pairs arranged around the circle are (a, a), (a, b), (b, b), (b, a), (a, c), (c, c), (c, b), (b, c), and (c, a). Every ordered pair in the set $\{a, b, c\}$ occurs once and only once. We shall demonstrate constructively the general existence of such arrangements of permutations.

THEOREM. *If n is a positive integer, and S is a finite set with $m \geqq 2$ elements, there exists a circle of m^n elements such that each n-tuple in S cccurs exactly once on (i.e., clockwise around) the circle.*

Before giving the proof, we present two definitions and a lemma. Throughout the remainder of the paper, all elements are tacitly assumed to belong to a finite set S with $m \geqq 2$ elements, and n is a positive integer.

DEFINITION. *A circle of elements is* **n-irredundant** *if no n-tuple occurs more than once on the circle.*

DEFINITION. *A circle of elements is* **n-balanced** *if whenever an n-tuple $(t_1, \cdots, t_n)$ occurs on the circle, so does each rotation $(t_i, \cdots, t_n, t_1, \cdots, t_{i-1})$, $1 < i \leqq n$, of $(t_1, \cdots, t_n)$.*

LEMMA. *Let C be an n-balanced circle of $p \geqq n$ elements such that whenever $(t_1, \cdots, t_n)$ occurs on C and $s \in S$, the n-tuple $(s, t_2, \cdots, t_n)$ also occurs on C. Then each n-tuple in S occurs on C.*

Proof of the Lemma. Suppose $(x_1, \cdots, x_n)$ is any n-tuple in S. Let $(y_1, \cdots, y_n)$ be an n-tuple which occurs on C. By hypothesis, $(x_1, y_2, \cdots, y_n)$ also occurs on C. If $n = 1$, the lemma is proven. Otherwise, note that $(y_2, \cdots, y_n, x_1)$ occurs on C since C is n-balanced. By a second application of the hypothesis, $(x_2, y_3, \cdots, y_n, x_1)$ also occurs on C. Each application of the hypothesis enables us to replace a y_i by an x_i, $1 \leqq i \leqq n$. The lemma is established by n such steps.

Proof of the theorem. Given an n-irredundant, n-balanced circle C with p elements, $n \leqq p < m^n$, we show how a larger n-irredundant, n-balanced circle D may be constructed. The existence of an n-irredundant, n-balanced circle with $\geqq m^n$ elements then follows from the existence of an n-irredundant, n-balanced circle with n elements. (One such circle is obtained as follows: Let $a, b \in S$ with $a \neq b$. Form the circle of n elements in which the element a occurs $n-1$ times and b occurs once.) Since there exist exactly m^n n-tuples in S, any n-irredundant circle with $\geqq m^n$ elements has exactly m^n elements, and every n-tuple occurs exactly once on it.

Now let C be given. By the lemma there exists an n-tuple $(x_1, \cdots, x_n)$ on C and an element $s \in S$ such that $(s, x_2, \cdots, x_n)$ is not on C. Define $u_1 = s$, and $u_i = x_i$ for $2 \leqq i \leqq n$. Let k be the smallest positive integer dividing n such that $u_i = u_{i-k}$ for $k < i \leqq n$. (Usually $k = n$. An example in which $k = 2$ and $n = 6$ is $(u_1, \cdots, u_n) = (a, b, a, b, a, b)$.)

Now construct the circle D from C by inserting the elements $u_1, \cdots, u_k$ into C just after the n-tuple $(x_1, \cdots, x_n)$. We shall present an example to show how this construction works. The reader may then complete the proof of the theorem himself by showing that the circle D in the proof is n-irredundant and n-balanced.

EXAMPLE. Let $S = \{a, b, c\}$ and let $n = 3$. As our starting point we choose the circle of 3 elements in which a, a, b occur in that order clockwise around the

circle. At each step we show our choice of (x_1, x_2, x_3) and s. Usually this choice is somewhat arbitrary. To make the construction clearer, at each step the new portion of the circle is underlined.

Circle	Step	(x_1, x_2, x_3)	s	k	$(u_1, \cdots, u_k)$
aab					
	1	aab	b	3	bab
$aab\underline{bab}$					
	2	aab	c	3	cab
$aab\underline{cab}bab$					
	3	abb	c	3	cbb
$aabcabb\underline{cbb}ab$					
	4	cbb	b	1	b
$aabcabbcbb\underline{b}ab$					
	5	baa	c	3	caa
$aa\underline{caa}bcabbcbbbab$					
	6	bba	c	3	cba
$aacaabcabbcbbba\underline{cba}b$					
	7	acb	c	3	ccb
$aacaabcabbcbbbacb\underline{ccb}ab$					
	8	caa	a	1	a
$aacaa\underline{a}bcabbcbbbacbccbab$					
	9	bcc	c	1	c
$aacaaabcabbcbbbacbcc\underline{c}bab$					
	10	aca	c	3	cca
$aaca\underline{cca}aabcabbcbbbacbcccbab$					

The purpose of k can be ascertained from Step 4. If bbb had been inserted instead of b, we would have obtained the circle $aabcabbcb\underline{bbb}bab$ which is not 3-irredundant.

Note that at each step in the construction, a new circle D is formed from an old circle C in such a manner that the n-tuples which occur around D are just the n-tuples which occur around C together with the new n-tuples $(u_1, \cdots, u_n)$, $(u_2, \cdots, u_n, u_1)$, etc.

A GROUP WHOSE SQUARES GENERATE A DICYCLIC GROUP

H. S. SUN, Fresno State College

A two-generator group $\langle a, b\rangle$ with defining relations: $a^{2m}=b^4=1$, $a^m=b^2$, $b^{-1}ab=a^{-1}$ is called a **dicyclic group** [1, p. 182]. G. A. Miller stated that no dicyclic group can be generated by the squares of any group [2, p. 152]. The following theorem shows that this statement is false:

THEOREM. *Let D be a dicyclic group of order $4m$ with defining relations:*

$a^{2m}=b^4=1$, $a^m=b^2$, $b^{-1}ab=a^{-1}$. *Then there exists a group G whose subgroup generated by its squares is isomorphic to D if and only if $t^2\equiv -1$ (mod $2m$) has a solution.*

Proof. Suppose G exists. We may assume that $m>2$, since, for $m=2$, D is not generated by the squares of any group [**3**, p. 194]. Identify D with the group generated by the squares in G. The $4m$ elements in D may be written as $1, a, a^2, \cdots, a^{2m-1}, b, ab, a^2b, \cdots, a^{2m-1}b$. Notice that each a^ib transforms a to its inverse. At least one of the a^ib must be a square in G, say, $a^ib=d^2$, for some i. Since $\langle a\rangle$ is a characteristic subgroup of D, it is normal in G; $d^{-1}ad=a^t$, but

$$a^{-1} = (a^ib)^{-1}a(a^ib) = d^{-1}(d^{-1}ad)d = d^{-1}a^td = a^{t^2}.$$

Hence $t^2\equiv -1$ (mod $2m$) must have a solution.

Conversely, if $t^2\equiv -1$ (mod $2m$) has a solution t_0, we define the group $G=\langle c, d\rangle$ by the relations: $c^{4m}=d^8=1$, $d^4=c^{2m}$ and $d^{-1}cd=c^{t_0}$. Clearly, $D=\langle c^2, d^2\rangle$ is a dicyclic group of order $4m$ and is generated by the squares in G.

The case $m=5$, $t=3$ shows that this situation actually occurs.

References

1. R. D. Carmichael, Introduction to the Theory of Groups of Finite Order, Dover, New York, 1956.
2. G. A. Miller, The Collected Works of George Abram Miller, Vol. 5, University of Illinois Press, 1959.
3. W. R. Scott, Group Theory, Prentice-Hall, Englewood Cliffs, N. J., 1964.
4. H. S. Sun, Embedding a Set into a Group, dissertation, University of New Brunswick, 1969.

CHARACTERIZING THE CIRCLE

WILLIAM KOENEN, Highland High School, St. Paul, Minn.

In the 1963 symposia on convexity, sponsored by the American Mathematical Society, Ludwig Danzer presented an elementary proof of a theorem of Besicovitch: *Assume $\mathcal{C}$ is a planar closed convex curve and no rectangle has exactly three vertices on $\mathcal{C}$; then $\mathcal{C}$ is a circle.* The author asks in that paper whether the theorem is true without the hypothesis of convexity.

The convexity restriction may indeed be deleted from the hypotheses because of the following theorem:

THEOREM. *Any simple closed curve in the plane (i.e., a compactum homeomorphic to a circle) is convex if it contains the fourth vertex of each rectangle whose remaining three vertices lie on the curve.*

The development of the main theorem requires an elementary result which apparently does not appear in the literature. Accordingly, this is handled first.

In a plane, if a line ℓ contains at least one point of a simple closed curve, and if the curve does not have points on both sides of ℓ, then ℓ is a **support line** for the

simple closed curve. The support lines help to describe those simple closed curves which are convex.

LEMMA. *A simple closed curve in a plane is convex if and only if along every line of support the contact points make a connected set.*

Proof. Exterior to a nonconvex simple closed curve, there must be a point between some two interior points, and therefore the set of rays from that exterior point through points of the curve has angular measure greater than π. There is also another exterior point far enough from the curve for the angular measure of the set of rays from that point through points of the curve to be arbitrarily near zero. Passing from one to the other of these positions, through the exterior of the curve, the measure of these rays varies continuously and thus assumes π as an intermediate value. The exterior point from which the set of rays through points of the curve has measure π is seen to be a point on a support line which is not a point of contact (it is an exterior point) but it lies between two points of contact. Each nonconvex curve has a support line whose contact points do not make a connected set. Convex curves, together with their interiors, make convex sets which must intersect support lines in convex (hence connected) sets. Support lines contain no interior points, so convex curves contact support lines in connected sets.

The main result also requires separation principles:

I. A Jordan arc joining a point on one side of a line to a point on the other side has at least one point in common with the line.

II. A Jordan arc joining a point interior to an angular region (i.e., the intersection of two open half-planes) to an exterior point has at least one point in common with the boundary.

III. (Jordan Curve Theorem.) A simple closed curve in a plane divides the remainder of the plane into two components: the exterior, having angular index zero, and the interior, having angular index ± 1. A Jordan arc joining an interior point to an exterior point has at least one point in common with the simple closed curve. [1]

IV. If A and B are two points of a simple closed curve, then the curve is the union of two Jordan arcs joining A to B. We call these arcs the **branches** of the curve.

V. If A, B, P, and Q are four distinct points of a simple closed curve $\mathcal{C}$ in a plane, and if P and Q are not on the same branch from A to B, then any two Jordan arcs inside $\mathcal{C}$, one joining A to B, the other joining P to Q, must have at least one point in common.

Now let $\mathcal{C}$ be any simple closed curve in the plane having a support line ℓ with two points of contact A and B. Assume further that there is no rectangle with exactly three vertices on $\mathcal{C}$. Let $\mathcal{S}$ be the semicircle from A to B on the same side of ℓ with $\mathcal{C}$. Evidently, $\mathcal{C}$ can have no points other than A or B in common

with $\mathcal{S}$, for in that event, $\mathcal{C}$ would contain exactly three vertices of a rectangle.

The portion of the closed upper half-plane bounded by ℓ inside (outside) the circle containing $\mathcal{S}$ we refer to as **inside (outside)** $\mathcal{S}$. The two branches of $\mathcal{C}$ connecting A and B might both lie inside $\mathcal{S}$ or at least one of them lies outside $\mathcal{S}$.

Assuming first that both branches of $\mathcal{C}$ lie inside $\mathcal{S}$, we see immediately that there is a point X of $\mathcal{C}$ not on ℓ. Let m be the ray from X normal to AX which meets ℓ, and let n be the semicircle with AX as diameter, on the opposite side of AX from m. See Fig. 1.

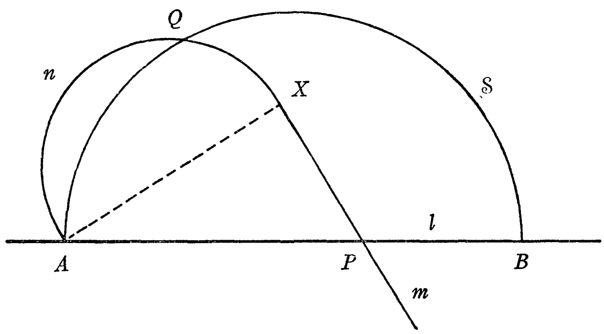


FIG. 1

Elementary geometric principles insure that m contains a point P between A and B and that n contains a point Q of $\mathcal{S}$ other than A or B. The branch of $\mathcal{C}$ not containing X must pass through the Jordan arc PQ at some point Y other than X, possibly $Y=P$. The points A, X, and Y are three vertices of a rectangle whose fourth vertex lies on the other side of the support line ℓ.

We have just shown that not both branches of $\mathcal{C}$ lie inside $\mathcal{S}$.

Let v be a line perpendicular to ℓ and separating A from B (see Fig. 2). Line v meets both branches of $\mathcal{C}$ at points we call X and Y, point X the one farther from ℓ. Since we have shown that at least one branch of $\mathcal{C}$ lies outside $\mathcal{S}$, we can be sure that X is outside $\mathcal{S}$.

The point Y must lie on ℓ between A and B. This is shown indirectly by noting that if Y is not on ℓ, then the line n through Y and parallel to ℓ separates X from A, so a point Z of the upper branch from A to B is on n. Three vertices X, Y, and Z of a rectangle lie on $\mathcal{C}$; therefore W, the fourth vertex, also lies on $\mathcal{C}$.

The ray p from X normal to AX and meeting ℓ, and the ray q from X normal to BX and meeting ℓ, make the boundary of an angular region whose interior contains A and whose exterior contains W. The curve $\mathcal{C}$ meets the boundary at least twice—we can be sure of some point $V\in\mathcal{C}$ other than X on p or q. If $V\in p$, then A, X, and V are three vertices of a rectangle whose fourth does not lie on $\mathcal{C}$. It is on the other side of the support line ℓ. Likewise, if $V\in q$, then B, X, and V lead to the same kind of impossibility.

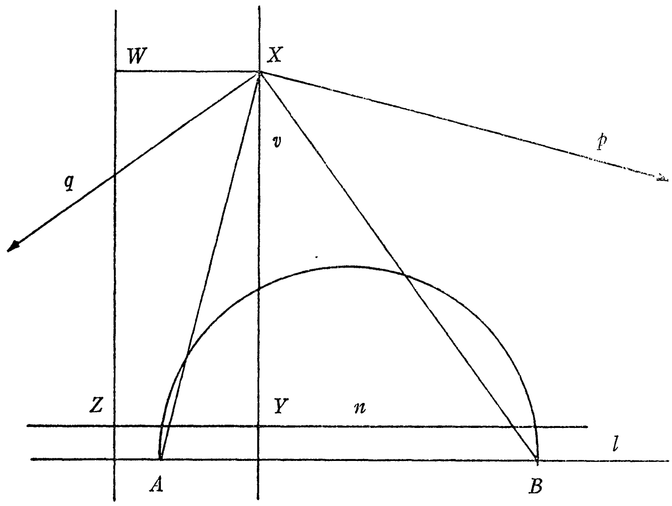


FIG. 2

Therefore Y must lie on ℓ between A and B. The lower branch of $\mathcal{C}$ contains only points of the segment AB.

In summary, every simple closed curve $\mathcal{C}$ with the property that no rectangle has exactly three vertices on $\mathcal{C}$, has the further property that any support line with two points of contact A and B must contact $\mathcal{C}$ along the entire segment AB. Such a curve $\mathcal{C}$ is convex, and, by Besicovitch's theorem, a circle.

References

1. P. S. Aleksandrov, Combinatorical Topology, vol. I, Graylock Press, New York, 1956.
2. Ludwig Danzer, A Characterization of the Circle, Proc. of Symposia in Pure Math., vol. VII, Convexity. A. M. S. 1963.

RATIONAL NUMBERS GENERATED BY TWO INTEGERS

G. A. HEUER, Concordia College

If a is a positive integer, the multiplicative group $\{a^n : n \in \mathbb{Z}\}$ generated by a is a discrete subset of the positive rationals. In contrast to this, the group $\{a^m b^n : m, n \in \mathbb{Z}\}$ generated by two positive integers (except in the trivial case where both are powers of a common integer) is dense in the positive real numbers. This interesting fact is not new [1], but a self-contained elementary proof has not, to my knowledge, appeared. I offer one here that uses no number theory beyond unique factorization of integers into powers of primes.

One of the consequences of this fact is that if a is a positive integer, not a power of 10, and $d_1, d_2, \cdots, d_n$ any finite sequence of decimal digits, then some integral power of a has $d_1 d_2 \cdots d_n$ as its initial sequence of digits. For if D is the integer with decimal representation $d_1 d_2 \cdots d_n$, we are asking that $10^k D \leqq a^j$

$<10^k(D+1)$ for some positive integers j and k; i.e., that $D \leqq a^j 10^{-k} < D+1$. There is an obvious generalization to an arbitrary base.

LEMMA 1. *Let a and b be positive integers. Then $\log_a b$ is rational if and only if a and b are integral powers of a common integer.*

Proof. This is where unique factorization is used; the proof is easy, and I leave it to the reader.

LEMMA 2. *If ξ is irrational, there are infinitely many pairs (h, k) of integers for which $0 < |\xi - h/k| < 1/k^2$. Thus for each m, there is such a pair with $k > m$.*

Proof. Although this result is readily accessible [2, p. 42] the proof is short and I include it. Let n be a positive integer. The $n+1$ numbers $0, \xi-[\xi], 2\xi-[2\xi], \cdots, n\xi-[n\xi]$ all lie in $[0, 1]$, so some subinterval $[j/n, (j+1)/n)$, $0 \leqq j < n$, contains two of them, $n_1\xi - [n_1\xi]$ and $n_2\xi - [n_2\xi]$. Let $k = n_1 - n_2$ and $h = [n_1\xi] - [n_2\xi]$, and note that $|k| \leqq n$. Then $0 < |k\xi - h| < 1/n$, whence $0 < |\xi - h/k| < 1/|nk| \leqq 1/k^2$. Thus, for each positive integer n, there is an integer pair (h, k) which simultaneously satisfies the two conditions $0 < |k\xi - h| < 1/n$ and $0 < |\xi - h/k| < 1/k^2$. It follows that there are infinitely many pairs (h, k) satisfying the latter condition.

THEOREM 1. *If ξ is irrational and $0 \leqq x < y \leqq 1$, then there is an integer m such that $x < m\xi - [m\xi] < y$.*

Proof. By Lemma 2, there are integers h, k with $|k| > 1/(y-x)$ such that $0 < |\xi - h/k| < 1/k^2$. Then $k\xi = h + \epsilon$, where $0 < |\epsilon| < 1/|k| < y - x$. Hence for some integer d, $x < d\epsilon < y$. Since $dk\xi = dh + d\epsilon$ and dh is an integer, the integer $m = dk$ has the desired property.

THEOREM 2. *If b and c are positive integers, not both integral powers of a single integer, then $\{b^m c^n : m, n \in \mathbf{Z}\}$ is dense in the positive real numbers.*

Proof. Let $0 < u < v$. Then $u < b^m c^n < v$ if and only if

$$(*) \qquad -n + \log_c u < m \log_c b < -n + \log_c v.$$

Let $x = \log_c u - [\log_c u]$ and $y = \log_c v - [\log_c v]$, and consider first the case $x < y$. By Lemma 1 and Theorem 1, there is an integer m such that the fractional part of $m \log_c b$ lies between x and y. Such m, together with $n = [\log_c u] - [m \log_c b]$, is readily seen to satisfy (*).

In the remaining case, since $\log_c u < \log_c v$, it follows that $x \geqq y$ implies $[\log_c u] < [\log_c v]$; thus $[\log_c u] + 1 \leqq [\log_c v] < \log_c v$. Then $x < 1 < \log_c v - [\log_c u]$. If m is an integer for which the fractional part of $m \log_c b$ lies between x and 1, and $n = [\log_c u] - [m \log_c b]$ again, one checks easily that (*) is again satisfied.

References

1. Leo Moser and Nathaniel Macon, On the distribution of first digits of powers, Scripta Math., 16 (1950) 290–292.

2. Ivan Niven, Irrational Numbers, MAA, Carus Monograph 11, 1956.

MAXIMAL HOMEOMORPHISMS

G. L. CANTRELL, Murray State University

1. Introduction. Maximal (minimal) sets with respect to various properties do not always exist. In practice, the existence of maximal (minimal) sets is usually established by use of the Hausdorff Maximal Principle or one of the many statements that are equivalent. In this setting the partial order is defined by set inclusion. The Hausdorff Maximal Principle allows the standard technique of considering the union of the elements in a maximal chain as a candidate for a maximal element, and the intersection of the elements in a maximal chain as a candidate for a minimal element. Theorem 2 gives an example of a situation in which there is a maximal set with the desired property, but the technique just mentioned will not always produce an acceptable candidate. This theorem is a corollary to theorems by J. de Groot [**1**, p. 486] and I. A. Vainstein [**2**, p. 34]. The proof presented here represents a somewhat different approach.

In the fourth part of this paper we investigate the occurrence of the maximal sets of Theorem 2 in a slightly more general setting. All topological spaces are taken to be Hausdorff, and we use $\overline{A}$ to denote topological closure of the set A.

2. A condition for a homeomorphism. The first theorem is used to prove Theorem 2.

THEOREM 1. *If X is a compact topological space and $f:X\to Y$ is a continuous function, then $M\subseteq X$ is a maximal set such that $f|M$ is a homeomorphism if and only if* (i) $f(\overline{M})=f(X)$ *and* (ii) $M=\{x\in\overline{M}:f^{-1}f(x)\cap\overline{M}=\{x\}\}$.

For the proof we need a result from [**3**, p. 12].

LEMMA 1. *If $A\subseteq X$ and $B=\{x\in\overline{A}:f^{-1}f(x)\cap\overline{A}=\{x\}\}$ is not empty, then $f|B$ is a homeomorphism.*

Proof of Theorem 1. Now suppose that M is a maximal set such that $f|M$ is a homeomorphism. If $w\in f(X)-f(\overline{M})$, then there is $x\in X-\overline{M}$ so that $f(x)=w$. It follows that $f|[M\cup\{x\}]$ is a homeomorphism. This proves that $f(\overline{M})=f(X)$.

Because M is dense in $\overline{M}$, we have $M\subseteq\{x\in\overline{M}:f^{-1}f(x)\cap\overline{M}=\{x\}\}$. By the maximality of M and Lemma 1,

$$M=\{x\in\overline{M}:f^{-1}f(x)\cap\overline{M}=\{x\}\}.$$

On the other hand, if $f(\overline{M})=f(X)$,

$$M=\{x\in\overline{M}:f^{-1}f(x)\cap\overline{M}=\{x\}\},$$

and $t\in X-M$, then there is a $y\in\overline{M}-\{t\}$ with $f(y)=f(t)$. Since M is dense in $\overline{M}$, the inverse of $f|[M\cup\{t\}]$ is not continuous at $f(t)$. Again by Lemma 1, M is a maximal set so that $f|M$ is a homeomorphism.

THEOREM 2. *If X is a compact metric space and $f:X\to Y$ is a continuous function, then there is $M\subseteq X$ that is a maximal set with respect to the property that $f|M$ is a homeomorphism.*

Proof. By using what has been termed the standard technique, one can show that there is a minimal closed set F in X such that $f(F)=f(X)$. We shall show that $M=\{x\in F: f^{-1}f(x)\cap F=\{x\}\}$ is dense in F. By Theorem 1, this is sufficient to complete the proof.

We take $f=f|F$ and use d to denote a metric on X. Now let $x\in F$ and $r>0$. We shall show that there is a $y\in N(x;r)=\{w\in F: d(w,x)<r\}$ so that $f^{-1}f(y)=\{y\}$. To simplify notation, set $B(x;r)=\{w\in F: d(w,x)\leqq r\}$.

If $\{x\}=F$, then it is clear that $M=F$. Otherwise, we may assume that $N(x;r)\neq F$. Because of the minimal nature of F, the set $F-f^{-1}f(F-N(x;r))$ is nonempty and open. Thus there is a $z\in F$ and $0<s<r/2$ so that

$$B(z;s)\subseteq F-f^{-1}f(F-N(x;r)).$$

It follows that $f^{-1}f(B(z;s))\subseteq N(x;r)$.

In an inductive manner one may select $\{z_n\}_{n=1}^{\infty}$ and $\{s_n\}_{n=1}^{\infty}$, with $s=s_1$ and and $z=z_1$ so that $f^{-1}f(B(z_{n+1};s_{n+1}))\subseteq N(z_n;s_n)$ and $0<s_{n+1}<s_n/2$, for $n=1,2,3,\cdots$. The sequence $\{z_n\}_{n=1}^{\infty}$ is Cauchy. Let $y=\lim z_n$. Then, $f^{-1}f(y)\subseteq N(z_n;s_n)$, where $n=1,2,3,\cdots$. Hence $f^{-1}f(y)=\{y\}\subseteq N(x;r)$.

3. An example. We show that consideration of a maximal chain may fail to produce an acceptable candidate for a maximal set with respect to the property of Theorem 2.

Let X be the closed unit interval and Y the unit circle in the complex plane, both with the relative topology of the complex plane. Denote by f the function whose values are determined by $f(x)=\exp 2\pi ix$, for $0\leqq x\leqq 1$. The spaces X and Y and the function f satisfy the hypothesis of Theorem 2. The collection

$$H=\{[0,a):0<a<1\}\cup\{[0,a]:0<a<1\}\cup\{0\}$$

is a maximal chain of sets, having the property that the function f restricted to each of these sets is a homeomorphism. Since the continuity of the inverse function fails at $f(0)$, the function f restricted to the union of these sets is not a homeomorphism.

4. On maximal homeomorphism spaces. We define the topological space X to be a **maximal homeomorphism (M.H.) space** if for each continuous function f with domain X, there is a maximal set $M\subseteq X$, so that $f|M$ is a homeomorphism.

By Theorem 2, compact metric spaces are M.H. spaces. On the other hand not all compact spaces are M.H. spaces, as the following example shows.

Let X be the space of nonconstant nondecreasing functions whose domain is $[0,1]$ and whose range is in $\{0,1\}$. Considered as a subspace of $\{0,1\}^{[0,1]}$ with the product topology, X is a compact space [4, p. 173]. The function K whose values are given by $K(f)=$ Lebesgue measure of $f^{-1}(1)$ is a continuous function from X onto $[0,1]$. In view of Theorem 1, should a maximal set N exist, then $\overline{N}=X$. However, for each $f\in X$, the set $K^{-1}K(f)$ has two elements. Consequently the space X in this example is not an M.H. space.

THEOREM 3. *The continuous image of a compact M.H. space is an M.H. space.*

Proof. If X is a compact M.H. space, $f: X \to f(X) = Y$ is continuous, and $g: Y \to Z$ is continuous, then there is a maximal set M in X such that $g \circ f | M$ is a homeomorphism. By Theorem 1 and the fact that f is continuous,

$$g(Y) = g(f(\overline{M})) \subseteq g(\overline{f(M)}) \subseteq g(Y).$$

Next we show that

$$f(M) \subseteq N = \{x \in \overline{f(M)} : g^{-1}g(x) \cap \overline{f(M)} = \{x\}\}.$$

If y and y_1 are distinct points in $\overline{f(M)}$ and $g(y) = g(y_1)$, then there are distinct points t and t_1 in $\overline{M}$ such that $f(t) = y$ and $f(t_1) = y_1$. Therefore, $g \circ f(t) = g \circ f(t_1)$. Consequently, $t \notin M$ and $f(t) = y \notin f(M)$. This proves that $f(M) \subseteq N \subseteq \overline{f(M)}$. Thus $\overline{N} = \overline{f(M)}$. This gives, $g(\overline{N}) = g(Y)$ and

$$N = \{x \in \overline{N} : g^{-1}g(x) \cap \overline{N} = \{x\}\}.$$

By Theorem 1, this completes the proof of the theorem.

In conclusion, we wish to point out that there are compact M.H. spaces that are not metrizable; the long line furnishes an example. This and other examples give rise to the conjecture that a compact, first axiom M.H. space is metrizable.

References

1. J. de Groot, Realizations under continuous mappings, Indag. Math., 12 (1950) 483–492.
2. I. A. Vainstein, On closed mappings, Učen. Zap. M. G. U., 155 (1955) 3–53.
3. K. Kuratowski, Topology, v. 2, Academic Press, New York and London, 1966.
4. R. A. Johnson. A compact non-metrizable space such that every closed subset is G-delta, this MONTHLY, 77 (1970) 172–176.

RESEARCH PROBLEMS

EDITED BY RICHARD GUY

In this Department the Monthly presents easily stated research problems dealing with notions ordinarily encountered in undergraduate mathematics. Each problem should be accompanied by relevant references (if any are known to the author) and by a brief description of known partial results. Manuscripts should be sent to Richard Guy, Department of Mathematics, Statistics, and Computing Science, The University of Calgary, Calgary 44, Alberta, Canada.

HOW MANY MAGIC CONFIGURATIONS ARE THERE?

U. S. R. MURTY, University of Waterloo

Let S be a finite set of noncollinear points in R^2. The maximal intersections of S with the lines of R^2 are called the **lines determined by** S. A set S together with the lines determined by S is called a **configuration.** We shall refer to a configuration by referring to the set of its points. A positive integral-valued function on S is said to be k-**magic** if the sum of its values on each line is k. A configuration

SOME EXAMPLES IN HOMOLOGY THEORY

M. H. HALL, Texas Tech University

1. Introduction. Some of the most striking and important recent results of algebraic topology have concerned the relationships between different categories, such as the categories of differential and piecewise linear manifolds, and the use of various functors to clarify these relationships. (See the paper by P. J. Hilton in [3], pages 1–22 for an excellent exposition of these results.) The details of the pertinent examples and theorems are, however, beyond the level of a typical first course in algebraic topology. The purpose of this note is to present an example of a reasonable category **C**, different from the category **Top** of all pairs, and a homology theory on **C**, distinct from singular theory, which detects, by means of a suitable probe space X, the distinction between **C** and **Top**. (We require a homology theory to satisfy all the Eilenberg-Steenrod axioms, including the dimension axiom.) This theory can be presented as soon as singular theory has been presented.

2. The Category. By **map** or **mapping,** we mean a continuous function. It is well known that closed maps (a function is **closed** if the image of each closed set is closed) are in certain situations more natural and useful than arbitrary maps, as they preserve topological properties which may be lost under arbitrary mappings. For example, if $f:X\rightarrow Y$ is a closed surjective map, then Y has the identification topology, and if X is normal or paracompact, then so is Y. (See [1], Theorem VI, 1.4, page 121, Theorem VII, 3.3, page 145, and Theorem VIII, 2.6, page 165.)

This suggests that one consider a category, all of whose maps are closed. In order to obtain an admissible category, in the sense of [2], one is forced to consider only pairs (X, A) with X a T_1-space and A closed in X. It is easy to verify that the category **C** of all such pairs and all closed maps of such pairs is indeed an admissible category for a homology theory. We observe that **C** contains all compact Hausdorff pairs and all maps of such pairs, and thus any homology theory on **C** must agree with singular theory on the full subcategory of compact C-W pairs. (See [4], Corollary 9.2, page 52.)

3. The Homology Theory. The coefficients of all homology theories in this note will be the integers and will not be expressly mentioned again in this section. We shall use the notation of [5]; in particular, $\Delta(X)$ will denote the singular chain complex of the space X. Since the restriction of a closed map to a closed subspace of the domain is a closed map, there is a subcomplex $\Delta C(X)$ of $\Delta(X)$, generated by the singular simplexes which are closed maps. The complex $\Delta C(X, A)$ of a pair is then defined in the usual manner. Passing to the homology of this chain complex, yields a graded group, which we shall denote by $HC(X, A)$. The verification of the Eilenberg-Steenrod axioms for HC is similar to the verification for singular theory. Since a singular simplex in a Hausdorff space is

necessarily a closed map, HC and singular homology must agree on all Hausdorff pairs (Y, B) with B closed in Y.

4. The Probe Space. Throughout this section, we shall let I denote the unit interval with the usual topology, $H(Y, B)$ will be the singular homology of the topological pair (Y, B), and Z will be the integers.

We wish to consider a space X obtained by equipping the unit interval with a new topology. In this topology, a basic neighborhood of 0 is to be a basic neighborhood in the usual topology and a basic neighborhood of some nonzero point x is to be the union of a usual basic neighborhood of x with a deleted usual basic neighborhood of 0. It is easy to verify that this specification of neighborhoods does yield a topology, coarser than that of I, and that the space X is T_1, but not Hausdorff. Furthermore, a subset having 0 as a limit point must have all of X as its closure, and it follows that the proper closed subsets of X are either closed subsets of I which do not contain 0 or closed subsets of I having 0 as an isolated point. This fact, together with a connectedness argument, yields a proof of the following lemma:

LEMMA. *The only path in X beginning at* 0 *which is a closed map is the constant path at* 0.

Now, the identity function $i: I \to X$ is clearly a map, and so X is path connected. Thus, $H_0(X) \simeq Z$ (see [5], Lemma 4.4.7, page 175). However, a straightforward computation (similar to Exercise 3B, Chapter VI, page 219 of [4]) using the preceding lemma shows that $HC_0(X) \simeq Z \oplus Z$. Thus H and HC are distinct homology theories.

Since $HC_0(X) \not\simeq Z$, it follows that X is not contractible in the category **C**. That is, there is no *closed* map $J: I \times X \to X$ such that $J(1, x) = x$ for all x in X and $J|\{0\} \times X$ is constant. However, the function $J: I \times X \to X$ defined by $J(t, x) = tx$ is a map. For, since a proper closed set in X can contain 0 only as an isolated point, it is immediate that the complement of the inverse image under J of a closed set is open. Thus X is contractible in the category **Top**. Hence, spaces which are homotopically equivalent in **Top** are not necessarily homotopically equivalent in **C**, and the functor HC may be used to detect this. (This is, of course, not intended as a claim that HC can discriminate all such occurrences; such power is unusual throughout algebraic topology.)

The author wishes to thank the referee for pointing out a serious defect in an earlier version of this note.

References

1. James Dugundji, Topology, Allyn and Bacon, Boston, 1966.

2. S. Eilenberg and N. Steenrod, Foundations of Algebraic Topology, Princeton Math. Ser., No. 15, Princeton Univ. Press, 1952.

3. P. J. Hilton, ed., Studies in Modern Topology, MAA Studies in Mathematics, Volume 5, Prentice-Hall, Englewood Cliffs, N. J., 1968.

4. Sze-Tsen Hu, Homology Theory, Holden-Day, San Francisco, 1966.

5. E. H. Spanier, Algebraic Topology, McGraw-Hill, New York, 1966.

SUBGROUPS OF A SUPERSOLUBLE GROUP

C. D. H. COOPER, Macquarie University, Sydney, Australia

A group G is **supersoluble** if there is a normal series $G=G_0>G_1> \cdots >G_n=1$ (each G_j normal in G) such that each quotient group G_j/G_{j+1} is cyclic. It is well known that a finite supersoluble group has subgroups of every possible order. The most direct proof of this is by induction on the order using the Schur-Zassenhaus theorem ([4], page 144). The proofs given by Deskins [2] and Bray [1], while not using this result, rely on P. Hall's theorem on the existence of a subgroup of order m in a finite soluble group of order mn, where m and n are coprime ([3], page 141). Even this is rather heavy machinery for such special groups as supersoluble groups. The following is an elementary proof of this result.

THEOREM. *If G is a finite supersoluble group of order n and $m|n$, then G has a subgroup of order m.*

Proof: We prove the theorem by induction on $n=|G|$. (Subgroups and factor groups of supersoluble groups are supersoluble ([3], page 158).) Let $1<K<L< \cdots$ be a chief series for G. Then $|K|=p$ and $|L|=pq$ for primes p, q (perhaps equal). Let Q be a Sylow q-subgroup of L.

CASE 1: $p|m$. G/K has a subgroup H/K of order m/p and so $|H|=m$.

CASE 2: *$p\nmid m$ and $mp<n$.* G/K has a subgroup H/K of order m and so $|H|=mp<n$. Thus H has a subgroup of order m.

CASE 3: *$p\nmid m$, $mp=n$ and $N_G(Q)<G$.* Then $q\neq p$ and $|Q|=q$. Now $G=N_G(Q)L$ ([4], page 129) and so $G/L\cong N_G(Q)/N_L(Q)$. Since $N_G(Q)<G$, $N_L(Q)<L$. Thus $N_L(Q)=Q$ and so $|N_G(Q)|=(mp/pq)q=m$.

CASE 4: *$p\nmid m$, $mp=n$ and $N_G(Q)=G$.* Thus Q is a normal subgroup of G and since $q|m$, G/Q has a subgroup H/Q of order m/q whence $|H|=m$.

References

1. H. G. Bray, A note on CLT groups, Pacific J. Math., 27 (1968) 229–231.
2. W. E. Deskins, A characterization of finite supersolvable groups, this MONTHLY, 75 (1968) 180–182.
3. M. Hall, The Theory of Groups, Macmillan, New York, 1959.
4. E. Schenkman, Group Theory, Van Nostrand, Princeton, N. J., 1965.

INDEFINITE CUT SETS FOR REAL FUNCTIONS

FOSTER BROOKS, Kent State University

The study of the properties of a real function $y=f(x)$ commonly involves the use of sets such as $\{x:f(x)<\ell\}$, (variously $\leqq$, $>$, $\geqq$), for arbitrary levels ℓ. The purpose of this note is to suggest a slightly more general class of sets that serve the same purposes somewhat more conveniently. Since sets such as $\{x:f(x)<\ell\}$ are determined uniquely by f and ℓ, we shall call them **definite cut sets,** the word "cut" arising from an analogy with Dedekind cuts. The sets proposed in

this note generally are not unique, being arbitrary over the set of x's where $f(x) = \ell$; they will be called **indefinite cut sets.** Their advantage comes mainly from the fact that in a limit process, functional values below a given level ℓ produce not only limits that are below ℓ, but also, generally, some of the limits that equal ℓ. Hence definite cut sets generally do not reproduce their own kind under a limit process, whereas indefinite ones do.

The terminology and notation we shall use are as follows:

DEFINITION. If $f(x)$ is a real function and ℓ is a real number, sets $A(f, \ell)$ and $B(f, \ell)$ are called respectively **lower** and **upper indefinite cut sets** for the domain of f at the level ℓ if

$$\{x: f(x) < \ell\} \subseteq A(f, \ell) \subseteq \{x: f(x) \leqq \ell\} \text{ and } \{x: f(x) > \ell\} \subseteq B(f, \ell) \subseteq \{x: f(x) \geqq \ell\}.$$

If two such sets are complementary in the domain, then the ordered pair $(A(f, \ell) \mid B(f, \ell))$ is called an **indefinite level cut** in the domain by the function f at the level ℓ.

It follows immediately that indefinite cut sets are monotone with respect to ℓ in the sense that if $\ell_1 < \ell_2$, then $A(f, \ell_1) \subseteq A(f, \ell_2)$ and $B(f, \ell_1) \supseteq B(f, \ell_2)$. Conversely, if a system of such sets is prescribed in any way to include at least one for each of an everywhere dense set of levels ℓ (e.g., all rational levels) subject only to monotoneity on ℓ in the above sense, then the function $f(x)$ is uniquely determined provided that $f(x)$ is permitted to take on the values $+\infty$ and $-\infty$ as well as finite real values. This follows since if x is given, then $f(x)$ must equal the greatest lower bound of the set of all ℓ's for which x belongs to a lower cut set at level ℓ. This includes $f(x) = +\infty$ if x belongs to no lower cut set, and $f(x) = -\infty$ if x belongs to all lower cut sets. It follows also that definite cut sets can be given in terms of indefinite ones as denumerable unions or intersections. For example, $\{x: f(x) < \ell\} = \bigcup_{k<\ell} A(f, k)$, where the union may be taken over any set of levels k below ℓ with ℓ as a limit.

Usual properties of functions can be expressed readily in the language of indefinite cut sets. For example, it is easily seen that a function is continuous if and only if at each level (or at each of an everywhere dense set of levels) there exist both lower and upper cut sets that are open sets (or closed sets). A function is lower (upper) semicontinuous if and only if at each such level lower (upper) indefinite cut sets exist that are closed (or upper (lower) ones that are open); it is measurable if and only if at each such level there exist indefinite cut sets that are measurable; etc. Also it is clear that at each level the closure $\mathrm{Cl}\, A(f, \ell)$ and the derived set $A'(f, \ell)$ of lower indefinite cut sets for $f(x)$ are lower indefinite cut sets respectively for the infimum function and the lower limit function for $f(x)$. Since both the closure and the derived set for an arbitrary set are closed, it follows that both the infimum function and the lower limit function are always lower semicontinuous, and similarly that the supremum and the upper limit functions are upper semicontinuous.

Two common theorems in elementary real variables that admit naturally the use of indefinite cut sets in their proofs are:

(1) *The set of points of discontinuity of a lower* (*upper*) *semicontinuous function*

is an exhaustible set (first category of Baire), and

(2) *The set of points of discontinuity of the limit function for a pointwise convergent sequence of continuous real functions is an exhaustible set.*

A proof for the second will be given; a quite similar proof holds for the first, which will be left as an exercise.

Preliminary to this proof we note that in connection with limits of sequences of functions, indefinite cut sets lead to an especially useful analogy with limits of sequences of sets. To see this, we let $f_n(x)$ denote a sequence of real functions and denote by $\liminf f_n(x)$, $\limsup f_n(x)$, and $\lim f_n(x)$ respectively the lower and upper limit functions and, where it exists, the limit function of the sequence. The lower and upper limit functions always exist at each point x, provided we allow them to take on the values $+\infty$ and $-\infty$ as well as all finite reals; if they are identical, their common value is the limit function. Also, if S_n is a sequence of sets, we define lower and upper limit sets in the usual way as

$$\liminf S_n = \bigcup_{k=1}^{\infty} \bigcap_{n=k}^{\infty} S_n \quad \text{and} \quad \limsup S_n = \bigcap_{k=1}^{\infty} \bigcup_{n=k}^{\infty} S_n.$$

If they are identical, their common value is the limit set of the sequence, denoted by $\lim S_n$. In this language we have the following result:

THEOREM 1. *If $f_n(x)$ is any sequence of real functions, then at each level ℓ we have*

$$\begin{aligned} B(\liminf f_n, \ell) &= \liminf B(f_n, \ell), \\ B(\limsup f_n, \ell) &= \limsup B(f_n, \ell), \\ A(\liminf f_n, \ell) &= \limsup A(f_n, \ell), \\ A(\limsup f_n, \ell) &= \liminf A(f_n, \ell), \end{aligned}$$

in the sense that no matter what indefinite cut sets are used in the right members of these equations, they will produce some particular indefinite cut sets as shown for the left members.

Proof. If ℓ is given, and if x is a point where $\liminf f_n(x)$ exceeds ℓ, then $f_n(x)$ must exceed ℓ for all but a finite number of n. Thus x belongs to all but a finite number of the sets $B(f_n, \ell)$, and hence to their lower limit. Conversely, if x is a point that belongs to all but a finite number of the sets $B(f_n, \ell)$, then $f_n(x) \geqq \ell$ for all but a finite number of n. This assures us that the lower limit function at that point has its value at least as great as ℓ. Hence $\liminf B(f_n, \ell)$ contains all points where the lower limit function exceeds ℓ, possibly some points where it equals ℓ, and no others, thus making it a lower indefinite cut set for the lower limit function. Similar arguments hold for the other cases.

We now proceed with the proof of the pointwise limit theorem. Let $f_n(x)$ be a sequence of continuous real functions that converges pointwise to the limit function $\ell(x)$. Clearly $\ell(x)$ is discontinuous at a given point ξ if and only if its supremum exceeds its infimum at ξ. This is equivalent to the existence of two rational numbers α and β, with $\alpha < \beta$, for which ξ belongs both to all $\mathrm{Cl}\, A(\ell(x), \alpha)$

and to all Cl $B(\ell(x), \beta)$. Hence the set D of all points of discontinuity of $\ell(f)$ is given by

$$D = \bigcup_{\alpha<\beta,\ \text{rat}} (\text{Cl } A(\ell(x), \alpha) \cap \text{Cl } B(\ell(x), \beta)), \tag{1}$$

where the union is taken over all rational α and β with $\alpha<\beta$, and where it is immaterial what indefinite cut sets are used on the right. This is a denumerable union of sets, each of which will be seen to be nowhere dense, hence it is exhaustible. For since the functions $f_n(x)$ are continuous, they have both lower and upper indefinite cut sets at all levels that are open sets. Hence by Theorem 1 there exist lower and upper indefinite cut sets at all levels for $\ell(x)$ that are denumerable intersections of denumerable unions of open sets, hence G_δ sets. Substitute these sets in equation (1); if any set in the union were dense in any interval I, both Cl $A(\ell(x), \alpha)$ and Cl $B(\ell(x), \beta)$, hence also $A(\ell(x), \alpha)$ and $B(\ell(x), \beta)$, would be dense in I. The latter, being G_δ sets, would then both be residual (complement of exhaustible) in I, so they would have points x in common. But this is impossible since $\ell(x) \leqq \alpha < \beta \leqq \ell(x)$ at such points.

If one tries to extend this argument to lower or upper limit functions for a nonconverging sequence of continuous functions, it fails because only one of the indefinite cut sets for an extreme limit function can be proved by Theorem 1 to be a G_δ; the other is a $G_{\delta\sigma}$. However, essentially the same argument does hold to prove that the set of points of lower (upper) semidiscontinuity of a lower (upper) limit function of a sequence of continuous functions is exhaustible. For if we denote, say, the upper limit function by $u(x)$, then the set D_u of points where $u(x)$ is upper semidiscontinuous is given by

$$D_u = \bigcup_{\alpha<\beta,\ \text{rat}} (A(u(x), \alpha) \cap \text{Cl } B(u(x), \beta)).$$

Here $B(u(x), \beta)$ may be taken to be a G_δ as before, but $A(u(x), \alpha)$ may not be; however since it itself appears, not its closure, it need not be. For now if Cl $B(u(x), \beta)$ is dense in any interval I, then $B(u(x), \beta)$ is dense and thus residual in I as before. This requires $A(u(x), \alpha)$ to be exhaustible in I, since otherwise it would have points in common with $B(u(x), \beta)$, which is impossible. Hence $A(u(x), \alpha) \cap \text{Cl } B(u(x), \beta)$ is exhaustible in every interval I in which it is dense. This clearly makes it exhaustible, since the part of it not contained in any interval where it is dense, is nowhere dense.

While the last argument uses the fact from Theorem 1 that the $B(u(x), \beta)$ can be taken as G_δ sets, which in turn depends upon the fact that the $B(f_n, \beta)$ can be taken as open sets, for this we do not need the full continuity of the functions of the original sequence, only lower semicontinuity. Hence, we have the stronger and somewhat unusual result that the upper (lower) limit function for a sequence of lower (upper) semicontinuous functions is upper (lower) semicontinuous everywhere, except for an exhaustible set.

While for simplicity of exposition only real functions of one real variable have been mentioned, it is clear that indefinite cut sets apply to more general functions; the only essential requirement is that the ranges be ordered sets.

CLOSED SUBGROUPS OF A LOCALLY COMPACT GROUP

D. H. ANDERSON, Southern Methodist University

Let G be a locally compact T_0 group written multiplicatively, and suppose u is a complex-valued regular Borel measure on G. If $u_t(E)=u(tE)$ for each E in the collection $\mathcal{B}$ of all Borel sets and $t\in G$, define $P_u=\{t\in G: u_t=u\}$ (called the **periodic group** of u).

LEMMA. *Let $u=v+iw=u_1-u_2+iu_3-iu_4$, where v and w are real-valued and u_1-u_2, u_3-u_4 are the Jordan decompositions of v, w respectively. If $y\in G$ is such that $u(yE)=u(E)$ for $E\in\mathcal{B}$, then $u_n(yE)=u_n(E)$ for $n=1,\cdots,4$.*

THEOREM. *The set P_u is a closed subgroup of G. Moreover, every closed subgroup of G is the periodic group of some measure.*

Proof. Part of the first portion of the proof is modeled after [**1**, p. 275]. Let t, $z\in P_u$. If $E\in\mathcal{B}$, then $z^{-1}E\in\mathcal{B}$ and

$$u((tz^{-1})E) = u(z^{-1}E) = u(z(z^{-1}E)) = u(E).$$

Thus $tz^{-1}\in P_u$ and hence P_u is a subgroup of G. Now let $x\in\overline{P}_u$ (the closure of P_u in G). Write $u=u_1-u_2+iu_3-iu_4$, where each u_n is nonnegative and regular. Consider u_1; to get $u_1(xE)=u_1(E)$, it is sufficient to prove that $u_1(xK)=u_1(K)$ for any compact $K\in\mathcal{B}$ since u_1 is regular. Let U be any open set containing the compact set xK. Choose an open neighborhood W of e (identity of G) such that $WxK\subset U$. Since $x\in\overline{P}_u$ and Wx is a neighborhood of x, there is a $y\in P_u$ such that $y\in Wx$. By the lemma,

$$u_1(K) = u_1(yK) \leqq u_1((Wx)K) \leqq u_1(U).$$

Thus $u_1(K)\leqq u_1(xK)$ since u_1 is regular. Now $x^{-1}\in\overline{P}_u$ because $\overline{P}_u$ is also a subgroup; the above result applied to x^{-1} and xK substituted for x and K respectively gives

$$u_1(xK) \leqq u_1(x^{-1}(xK)) = u_1(K).$$

Hence $u_1(xK)=u_1(K)$ for any compact $K\in\mathcal{B}$, so $u_1(xE)=u_1(E)$. Similarly, $u_n(xE)=u_n(E)$ for $n=2, 3, 4$. Thus $x\in P_u$ which completes the proof that P_u is closed.

We now show that any closed subgroup H of G arises in this fashion. There exists a Haar measure m_H on H since H is a locally compact Hausdorff space. Define $v(E)=m_H(E\cap H)$ for G-Borel sets. Then

$$v_t(E) = m_H(tE\cap H) = m_H(t(E\cap H)) = v(E)$$

if $t\in H$; therefore $H\subset P_v$. Now suppose $t\notin H$. Since H is closed in G, there is an open set V in G such that $t\in V$ and $V\cap H=\varnothing$. Select E_0 to be an open set containing e such that $tE_0\subset V$. Then $m_H(tE_0\cap H)=0$, yet $m_H(E_0\cap H)>0$ since $E_0\cap H$ is open in H. Thus there is an $E_0\in\mathcal{B}$ for which $v_t(E_0)\neq v(E_0)$; it follows that $t\notin P_v$ and so $P_v\subset H$.

REMARK. If $P_u = G$, then u is left translation invariant on G and u is a complex constant multiple of Haar measure on G.

Reference

1. P. R. Halmos, Measure Theory, Van Nostrand, Princeton, N. J., 1950.

MATHEMATICAL EDUCATION

EDITED BY J. G. HARVEY AND M. W. POWNALL

Material for this Department should be sent to either of the editors: J. G. Harvey, Department of Mathematics, University of Wisconsin, Madison, WI 53706; M. W. Pownall, Department of Mathematics, Colgate University, Hamilton, NY 13346.

A STUDY OF SECONDARY MATHEMATICS TEACHERS: WHAT INFLUENCES THEM TO LEAVE THE PROFESSION?

MARGARET H. KNEITZ, Dominican College

In a recent study of secondary mathematics teacher dropouts in Texas conducted through the University of Houston, participants were asked to evaluate their undergraduate education in mathematics. They were also asked to comment on their experiences, both in college and in teaching, and to make suggestions for improvements. Although the study considered many other factors which might influence mathematics teacher retention or nonretention, only information obtained from the participants concerning their undergraduate preparation and early teaching experiences will be considered for this article.

The Population. The sample population was made up of 1964 and 1965 graduates who were certified to teach mathematics in secondary school, and student teachers who were doing their practice teaching in mathematics during the fall semester of 1968. Secondary school was defined to be grades six through twelve in a public or private school. The names and addresses of these persons were obtained from nineteen colleges and universities in Texas who had graduated a minimum of ten certified, prospective mathematics teachers during the school years of 1963–1964 and 1964–1965. The names of these colleges and universities were obtained from the Texas Education Agency's report of *Teacher Supply and Demand in Texas.* A total of 326 persons agreed to participate in the study and were sent a survey instrument through the mail. These survey instruments were mailed during the last week of October and through November, 1968. Returns on the majority of the survey instruments were received within three weeks. Fewer than thirty failed to return their survey instrument in this time, and follow-up letters and additional instruments were sent to these persons during the month of December, 1968. By January 1969, all but thirteen of the instruments had been returned, and collection of the data ended. Of the 326 persons who had agreed to participate in the study, 316 returned their completed survey instrument; a return of 96.05 percent. From this 316, four groups were formed which consisted of the following:

1. Student Teachers—students who were practice teaching in mathematics during the fall semester, 1968; a total of 63.
2. Never—graduates who had not taught since graduation; a total of 23.
3. Teaching—graduates who were currently teaching in a public or private secondary school; a total of 138.
4. Taught—graduates who had taught, but were not currently teaching; a total of 92.

TABLE I

Items concerned with undergraduate preparation

Please circle the appropriate number	Extremely well		Adequately		Not at all
I feel that my undergraduate mathematics courses prepared me for teaching mathematics in the secondary school.	5	4	3	2	1
I feel that my undergraduate mathematics methods course was beneficial. (If applicable)	5	4	3	2	1
I feel that my student teaching experience prepared me for teaching.	5	4	3	2	1

Please check as many as are appropriate in each set.

Mathematics courses I taught during student teaching were:

____6th, 7th, or 8th grade mathematics.
____related mathematics.
____algebra I or II.
____trigonometry and analysis.
____other. (please specify)

____

Mathematics courses I am teaching now are:

____6th, 7th, or 8th grade mathematics.
____related mathematics.
____algebra I or II.
____trigonometry and analysis.
____other. (please specify)

____

Analysis and Interpretation of Data. The items shown in Table I were concerned with undergraduate mathematics preparation and appeared in this form on the survey instrument. These items represent only part of the original survey instrument, but are the items which are pertinent to this segment of the analysis.

The data obtained from the sample population in regard to their undergraduate preparation (refer to Table II) were as follows:

1. Of the 316 participants, 315 responded to the item concerning their undergraduate mathematics preparation. Of these, 84.5 percent felt that their mathematics courses had prepared them adequately or better, and only 5.4 percent felt that their mathematics courses had not prepared them at all.
2. Not all of the participants had done student teaching, but of the 304 who had, 89.5 percent felt that this experience had prepared them for teaching adequately or better, and only 3.3 percent felt their student teaching experience had not prepared them at all.

TABLE II

Undergraduate preparation of the sample population

		Not at all 1	2	Ade- quately 3	4	Extremely well 5	Total
Mathematics	N	17	32	103	90	73	315
	%	5.4	10.1	32.7	28.6	23.2	
Mathematics Methods	N	25	38	72	45	42	222
	%	11.2	17.2	32.4	20.3	18.9	
Student Teaching	N	10	22	61	99	112	304
	%	3.3	7.2	20.1	32.6	36.8	

3. Only 222 out of the sample population of 316 had taken a mathematics methods course, and 71.6 percent of these 222 felt it had prepared them adequately or better.

TABLE III

Comparison by groups of the adequacy of mathematics methods courses

		Not at all 1	2	Ade- quately 3	4	Extremely well 5	Total
Student Teachers	N	4	8	9	6	5	32
	%	12.5	25.0	28.1	18.8	15.5	
Never	N	3	2	6	5	0	16
	%	18.7	12.5	37.5	31.3	0.0	
Teaching	N	9	20	36	21	23	109
	%	8.3	18.4	33.0	19.2	21.1	
Taught	N	9	8	21	13	14	65
	%	13.8	12.3	32.4	20.0	21.5	
Total		25	38	72	45	42	222

When these responses were considered in relation to each of the four groups (refer to Table III), the following information arose:

1. Out of the 32 respondents who were student teachers and who had taken a mathematics methods course, 37.5 percent felt that the methods course had been of little or no value at all to them. The remaining 62.5 percent felt it had been of adequate value or more.
2. Sixteen of the 23 in the Never Group had taken a methods course in

mathematics; 31.2 percent felt it was of little or no value, 37.5 percent felt it was adequate, 31.3 percent felt it had prepared them well, and no one felt it had prepared them extremely well.

3. The teaching group had 109 out of 138 responding; 26.7 percent felt it had prepared them inadequately or not at all, 33.0 percent felt it had been adequate, and 40.3 percent felt it had prepared them well or extremely well.

It is interesting to note that, while 85.1 percent of the participating 1964 and 1965 graduates had taken a mathematics methods course, only 51.6 percent of the Student Teaching Group had taken a mathematics methods course. Many of the participants wrote that they would have taken a mathematics methods course if one had been offered. The major criticisms made by the participants of the methods course were:

1. A methods course in mathematics was only offered if the demand were large enough to make a full class. Therefore, a student could not depend upon being able to take it during any given semester. Sometimes several semesters passed before a full class developed. By this time a student may well have graduated without it since such a course is usually taken during the latter part of the junior year or during the senior year. This was the reason stated by many of the participants for not having taken a mathematics methods course.
2. The participants who had not had a mathematics methods course but who had taken a general methods course felt that the course was not beneficial. Since it was often offered simultaneously with other subjects such as English, history, or art, the general methods course was usually centered around the preparation of lesson plans and the use of audio-visual equipment. No attention could be given to a variety of methods for presenting the many mathematics concepts to a roomful of students at different achievement levels.

In analyzing the questions concerning the mathematics courses taught during student teaching and those being presently taught, it was observed that only seven out of the 293 who had taken student teaching, or who were presently doing student teaching, had taught, or were teaching, related mathematics. The most frequently taught courses during student teaching were algebra (sixty-eight), geometry (twenty-three), algebra with geometry (twenty-nine), and algebra with trigonometry (fourteen).

In contrast, those in the Taught Group who had taught related mathematics during their last teaching assignment numbered twenty-nine out of ninety-two. The most frequently given courses as their present teaching assignment for those in the Teaching Group were sixth-, seventh-, or eighth-grade mathematics (twenty-nine), geometry (fourteen), and algebra (fourteen).

The major complaints of the 1964 and 1965 graduates in relation to their first year of teaching were as follows:

1. They were not prepared for the negative factors of teaching such as discipline problems, the extent of preparation time, and the lack of available materials and useable texts for slow learners.
2. Their assignment consisted of all, or a majority of, slow learners.
3. Their schedule frequently consisted of three or more preparations.
4. There was a great lack of administrative leadership and administrative support in the areas of discipline and the upholding of school policies.

Although low salaries were also included in their complaints, it was not, contrary to common opinion, the major problem.

Questions. What solutions can be found to satisfy such complaints? How can university teachers, school administrators, and experienced teachers assist new mathematics teachers and help make the first year or two of teaching more pleasant? Should mathematics methods courses place more emphasis on implementing the adopted textbooks, preparing learning situations which will help eliminate discipline problems, and the preparation of materials and methods of teaching the slow learners?

Can school administrators be influenced to see the need for assigning new mathematics teachers to a limited number of preparations, with an assignment of something other than a totality of slow learners?

The need to find a workable solution to these problems is readily apparent. The shortage of certified mathematics teachers has been critical for a number of years. Although the number of persons prepared to teach mathematics in the United States has increased almost 100 percent since 1950, the supply of newly produced, active candidates prepared to teach in high school has filled only 59 percent of the demand.

The turnover of teachers has exceeded 10 percent of the total profession every year, and fewer than three out of four people prepared to teach actually entered the classroom. Even more disturbing is the fact that less than 10 percent of this three-fourths will be found in the classroom in ten years. No statistics on the dropout rate of mathematics teachers could be found previous to this study.

It was disturbing to find that 115 out of the 253 graduates, 45.45 percent, were not teaching. Although they had prepared themselves for the teaching profession, 9.09 percent never entered the classroom. Of those who taught, 11.85 percent taught only one year, 10.67 percent taught only two years, and 13.83 percent taught three or more years before dropping out of the profession.

When the above questions have been carefully considered and action has been taken to improve the present situation, the yearly dropout rate of mathematics teachers will surely decrease.

References

National Education Association, Research Division, *Teacher Supply and Demand in Public Schools, 1965.* Washington: National Education Association, June, 1965.

R. H. Nelson, and Michael Thompson, Why Teachers Quit, *The Clearing House*, 37 (April, 1963) 467.

MATHEMATICS CURRICULA FOR DEVELOPING COUNTRIES

M. A. B. DEAKIN, Papua and New Guinea Institute of Technology

Recent articles in the MONTHLY (e.g., those by Kline [1], and Klamkin [2]) have criticized aspects of the syllabus revisions generally referred to as "the New Math." Such arguments have also been advanced elsewhere, by (e.g.) Feynman [3] and Hamming [4]. A recurrent theme of these articles is the discrepancy between a syllabus that stresses the mathematics required for research in an area that we in the British Commonwealth would designate as "Pure Mathematics," and the economic realities that dictate a need for proficiency in "Applied Mathematics."

One sidelight to this controversy deserves some publicity. American mathematics syllabuses are exported, particularly to developing countries. In some cases this is a conscious process, in that American consultants are called in, or in that a university in a developing country may enter a formal relationship with an American college or university. In other cases, the export is unconscious. An overworked and underqualified administrator in an emerging nation feels impelled to "keep up with modern developments overseas," and copies almost verbatim a course designed for (say) high schools in upper New York State.

Common to all but a very few mathematics departments is the fact that the vast majority of their teaching is service teaching. If this is true at the college level, it is even truer in the high schools. If it is true of the developed countries, it is even truer of those which are as yet underdeveloped.

The bulk of all mathematics education is not geared to the production of professional mathematicians, but rather aims to educate an intelligent user of mathematics in some field of its application. An emerging nation in particular needs to take cognizance of this fact. For such a country, manpower is the prime national resource—its wastage is intolerable. It therefore behooves those of us concerned with mathematical education in the third world to look long and hard at the suitability of our courses in terms of national needs and priorities.

The territory of Papua and New Guinea is, outside the well-publicized area of Southern Africa, the largest and most populous of those nations still described as "colonial." It comprises an administrative union of the Australian Territory of Papua and the U.N. Trust Territory of New Guinea. It is anticipated that the country will become independent under the name "Niugini," which will be used throughout the remainder of this article.

Our own institute is responsible for all professional training in the fields of engineering (including surveying), architecture, and commerce throughout the entire country. A separate university in the capital, Port Moresby, caters for other faculties, and, in particular, offers degrees with a mathematics major (we do not).

Those who graduate there in mathematics (there are very few of them) will mostly find employment as secondary teachers or government statisticians. The vast majority of those who study mathematics do so while pursuing courses in

engineering, science, or other fields. All these students are users of mathematical know-how—as indeed will be the mathematics majors themselves.

Our own institute teaches entirely to the nonspecialist. The high schools of Niugini are in the same position, essentially, as well below a quarter of one per cent of their graduates even seek to major in mathematics at the university in Port Moresby.

It is therefore disconcerting to administrators of tertiary mathematics here to find that the secondary schools have adopted a rather faddish "New Math" syllabus, loaded with set theory and its associated jargon. Actually, little set theory is really learned. The bulk of the pupil's effort is expended in the memorization of a specialist vocabulary (in at best his second, often his sixth, language).

Specific educational difficulties in this approach have been well documented by King [5] and myself [6]. The temptation is not unique to Niugini. Many African countries have adopted similarly ill-considered approaches to mathematics curricula (see, in particular, the Entebbe Mathematics Series [7]).

The unsuitability of these courses at the secondary level is evident to those of us charged with the instruction of engineers for a country with few roads and no railways. The grassroots local government engineer in Niugini will never need set theory or elementary topology. He will require a facility in calculation (in areas unequipped with electrical power), a knowledge of statistical data handling, and a basic grasp of a core of mathematical technique centered around manipulative algebra, geometry, trigonometry, calculus, and numerical methods. He need never have seen a formal proof in his life; he will never encounter a pathological function. I, for one, see no need to teach either concept.

Apart from these considerations, there is the real difficulty of implementing a "New Math" syllabus in a country of 2.5 million that has employed between one and nine qualified secondary mathematics teachers over the past decade (data from McKay [8]).

It provides not only solace, but genuine power to the arms of embattled math administrators such as myself, to see the effects of current debate in North America, e.g. [1, 2, 3, 4]. We are often hard put to make our point in view of the pressure that has been exerted on behalf of the "New Math."

It would be not only churlish, but absurd, in any sense to "blame" U. S. mathematicians for the overseas effects of syllabuses developed for home consumption. I would, however, draw the Association's attention to this second order effect, in the hope that its undesirable results might be minimized. In particular, I should like to urge those who become involved in consultation projects in education in the third world to bear in mind the aims of mathematics teaching in these countries. This may put heavy demands on the consultant, but unless these be met, his advice may well be useless or worse.

As a final aside, it occurs to me to ask whether some of these difficulties may not have already been encountered within the U. S. itself, particularly in those colleges whose student body is predominantly black. It could be that the syllabuses currently followed in emerging nations reflect not simply the trends of America, but more specifically, of White America.

References

1. M. Kline, Logic versus pedagogy, this MONTHLY, 77 (1970) 264–282.

2. M. S. Klamkin, On the ideal role of an industrial mathematician and its educational implications, this MONTHLY, 78 (1971), 53–76.

3. R. P. Feynman, New textbooks for the "New Mathematics," Engineering and Science, 28 (1965) 9–15.

4. R. W. Hamming, Numerical analysis versus mathematics, Science, 148 (1965), 473–475.

5. B. W. King, Objectives in mathematics: Why teach set theory?, Papua New Guinea J. Educ., 6 (5) (1970) 54–66.

6. M. A. B. Deakin, Mathematics curricula: Revision versus reform, Papua New Guinea J. Educ. (to appear).

7. Entebbe Mathematics Series, Educational Development Center, Newton, Mass.

8. M. H. McKay, What is new in mathematics?, Inaugural Lecture, University of Papua and New Guinea, 1968.

PROBLEMS AND SOLUTIONS

EDITED BY EMORY P. STARKE

All problems (both elementary and advanced) proposed for inclusion in this Department should be sent to E. P. Starke, 1000 Kensington Ave., Plainfield, NJ 07060. Proposers of problems are urged to enclose any solutions or information that will assist the editors. Ordinarily, problems in well-known textbooks and results in generally accessible sources are not appropriate for this Department. No solutions (except those accompanying proposals) should be sent to Professor Starke.

ELEMENTARY PROBLEMS

Solutions of Elementary Problems should be sent to Problems Group, Mathematics Department, University of Maine, Orono, ME 04473. To facilitate their consideration, solutions of Elementary Problems in this issue should be typed (with double spacing) and should be mailed before February 29, 1972. Contributors (in the United States) who desire acknowledgment of receipt of their solutions are asked to enclose self-addressed stamped postcards.

E 2319. *Proposed by Thomas Hern, Bowling Green State University*

If z_1 and z_2 are complex numbers with $0<|z_1|\leqq 1$ and $0<|z_2|\leqq 1$, show that $|z_1-z_2|\leqq|\log z_1-\log z_2|$.

E 2320. *Proposed by Erwin Just, Bronx Community College*

Let

$$\left\{\frac{a_1}{b_1},\frac{a_2}{b_2},\ldots,\frac{a_n}{b_n}\right\}$$

consist of n rational numbers in which the a_i and b_i are integers, and $(n, \prod_{i=1}^{n} b_i) = 1$. Prove that there exist positive integers k and m such that the numerator of the fraction determined by $\sum_{i=k}^{m} a_i/b_i$ is divisible by n.

E 2321. *Proposed by Michael Skalsky, Southern Illinois University*

Show that

$$\sum_{n=1}^{\infty} (nxe^{-x})^n/n! = x(1-x)^{-1}.$$

E 2322. *Proposed by Harry Lass and Peter Gottlieb, Jet Propulsion Laboratory, California Institute of Technology*

Let $A_1, \cdots, A_n$ be finite sets, each with the same number of elements, and let $S = \bigcup_{j=1}^{n} A_j$. Suppose that for some fixed k with $1 \leqq k \leqq n$, every union of k of the sets is S and every union of less than k of the sets is a proper subset of S. Determine in terms of n and k (1) the minimum number of elements in S; (2) the number of elements in each A_j when the number of elements in S is minimal; and (3) the number of elements common to any j of the subsets when S is minimal.

E 2323. *Proposed by Anders Bager, Hjørring, Denmark*

Characterize those triangles for which $\sqrt{3}+5\sum \cot A \geqq 3\sum \csc A$. (Here $\sum f(A)$ is taken to mean $f(A)+f(B)+f(C)$.)

E 2324. *Proposed by Frank Dapkus, Seton Hall University*

What is the probability that the length of a chord randomly drawn in an ellipse will not exceed the length of the minor axis? (By "randomly drawn chords" we mean those with midpoints uniformly distributed throughout the ellipse.)

SOLUTIONS OF ELEMENTARY PROBLEMS

A Relation Existing on Many Finite Sets

E 2239 [1970, 523; 1971, 409]. *Proposed by G. Sabbagh, Paris, France*

Let R be a binary relation on a nonvoid set E such that

(1) For no $x \in E$ is xRx true.

(2) For each pair (x, y) of distinct elements of E one and only one of the following relations holds: xRy, yRx.

(3) R is dense, which means: If xRy, then there is a $z \in E$ such that xRz and zRy.

Must E be infinite?

III. *Comment by D. Ž. Djoković, University of Waterloo, Canada.* Reference is to II by David Singmaster [1970, 410]. The only projective plane that can be obtained in this way is the plane with seven points. Indeed, assume that we have a projective plane by taking E to be the set of points and L_x for $x \in E$ to be the

lines. Assume also that this plane has more than seven points. Then there are at least four points on every line. Let $x \in E$ and choose $y \in E$ so that $y \neq x$ and $y \notin L_x$. By condition (2) of the problem we have $x \in L_y$. Let L be the line incident with x and y. Then we can choose $z \in L$ such that $z \neq x$, $z \neq y$ and $z \notin L_x$. By condition (2) again we have $x \in L_z$. By condition (1) $y \notin L_y$. Therefore $z \notin L_y$. Hence $y \in L_z$. Thus $L_z = L$ and we get $z \in L_z$ which contradicts (1).

Sets of Natural Numbers with Equal Sum and Product

E 2262 [1970, 1008]. *Proposed by G. J. Simmons and D. B. Rawlinson, Sandia Laboratories, Albuquerque, N. M.*

For every $k(>1)$ there is a set of k positive integers whose sum and product are equal. For $k=2, 3, 4, 6, 24$, the set is unique. Is it unique for any other k?

Solution by E. P. Starke, Plainfield, N. J. Call a set of k (not necessarily distinct) natural numbers whose sum and product are equal a *k-satisfactory* set. For any $k \geqq 2$, the set consisting of k and 2 and $k-2$ ones is a k-satisfactory set. For a and b and $k-2$ ones to form a k-satisfactory set it is necessary and sufficient that $(a-1)(b-1)=k-1$, so that whenever $k-1$ is composite, there will exist at least two k-satisfactory sets. Similarly the set consisting of a, b, 2, and $k-3$ ones will be a k-satisfactory set whenever $(2a-1)(2b-1)=2k-1$.

It follows then that a necessary condition that there be a unique k-satisfactory set is that both $k-1$ and $2k-1$ be prime. (For $k>6$ we must have then that k is a multiple of 6 whose last digit is neither 6 nor 8.) This eliminates from consideration all k through 232 except for the following: 2, 3, 4, 6, 12, 24, 30, 42, 54, 84, 90, 114, 132, 174, 180, and 192. Nine of these have second satisfactory sets as we now show (only the non-unit members are listed):

k	second k-satisfactory set
12	(2, 2, 2, 2)
30	(2, 2, 3, 3)
42	(2, 2, 2, 2, 3)
54	(2, 2, 2, 8)
84	(3, 4, 8)
90	(2, 2, 5, 5)
132	(2, 2, 6, 6)
180	(2, 2, 2, 26)
192	(2, 2, 2, 2, 13)

Thus we have eliminated all k through 232 except for 2, 3, 4, 6, 24, 114, and 174; for these the satisfactory set is unique. The calculations involved are not laborious, but to determine the situation for larger values of k, a computer would be useful.

Also solved by W. D. Bouwsma, C. Gardner, M. Hirschhorn (Scotland), L. E. Mattics, Norman Miller, Charles Wexler, W. G. Wild, and A. Zujus.

Editorial Comment. Wexler points out that for $a>1$, $b>1$, $c>1$, and $k-3$ ones, then $k-3=c(ab-1)-a-b$, so whenever $k\equiv 3-a-b \pmod{ab-1}$, then there is a second solution. Similarly for $a>1$, $b>1$, $c>1$, $d>1$ and $k-4$ ones, the condition is $k\equiv 4-a-b-c \pmod{abc-1}$. Using $a=2$ or 3, and $b=2, 3, 4$, or 6 in the first congruence; and $a=b=2$, and $c=2$ or 3 in the second congruence eliminates all $k<200$ except for 2, 3, 4, 6, 24, 114, and 174.

Bouwsma also finds that $k=444$ has a unique satisfactory set, "completing the list of all solutions not exceeding $k=1000$."

With the aid of a computer, this reviewer finds that 2, 3, 4, 6, 24, 114, 174, and 444 are the only solutions not exceeding 10,000. Now only when $k\geqq 2^n-n$ can k-satisfactory sets contain as many as n integers greater than one; it was found by computer that only seven values of k up to 4158 had satisfactory sets requiring five non-unit numbers, and only 32 others needed four such non-unit numbers. None required more than five. It is quite conceivable that there are no solutions to this problem other than the eight listed above.

The proposers remark that the problem arose in connection with computing the maximal cyclic subgroups of the symmetric group S_n.

Matching Balls Drawn from an Urn

E 2263 [1970, 1008; 1971, 404]. *Proposed by Bernard McCabe, Bell Comm. Inc., Washington, D. C.*

Suppose an urn contains m balls, each a different color. An observer draws a ball at random, records its color, and replaces it in the urn. He repeats this procedure until some color reappears for the first time. Show that the expected number of drawings is

$$E(m) = \sum_{k=1}^{m} \frac{k(k+1)m!}{(m-k)!\,m^{k+1}},$$

and determine the leading term in the asymptotic expansion.

I. *Solution by B. C. Arnold, Iowa State University.* Let X be the random variable equal to the number of drawings until some color reappears for the first time. Then, for any nonnegative integer k, $P(X>k)=P$ (the first k balls are all of different colors) $=(m)_k/m^k$, where $(m)_k=m(m-1)\cdots(m-k+1)$, and $(m)_0=1$. Thus, for $k=1, 2, \cdots, m+1$,

$$P(X=k) = P(X>k-1) - P(X>k)$$
$$= \frac{(m)_{k-1}}{(k-1)!} - \frac{(m)_k}{k!} = \frac{(k-1)m!}{(m-k+1)!\,m^k}.$$

(Obviously $P(X=k)=0$ if $k>m+1$.) Consequently

$$E(X) = \sum_{k=1}^{m+1} kP(X=k) = \sum_{k=1}^{m+1} \frac{k(k-1)m!}{(m-k+1)!\,m^k} = \sum_{j=0}^{m} \frac{j(j+1)m!}{(m-j)!\,m^{j+1}},$$

which is the desired expression for the mean, since the term for $j=0$ is 0.

To investigate the asymptotic behavior of $E(X)$ it is convenient to recall that if X is a nonnegative integer-valued random variable, then $E(X)=\sum_{k=0}^{\infty} P(X>k)$. In the present case, this implies that

$$E(X) = \sum_{k=0}^{m} \frac{(m)_k}{m^k} = \sum_{k=0}^{m} \frac{m!}{(m-k)!m^k}$$
$$= \sum_{j=0}^{m} \frac{m!}{j!m^{m-j}} = \frac{m!e^m}{m^m}\left[\sum_{j=0}^{m} \frac{e^{-m}m^j}{j!}\right].$$

We note that the term in square brackets is $P(Y \leqq m)$, where Y is a Poisson random variable with mean m. [*Ed. note*: This provides a convenient method for computing $E(X)$, since cumulative Poisson tables are readily available. It is well known that the determination can be made also from tables of the incomplete gamma function or of the χ^2 distribution.] But $P(Y \leqq m) = P(Y-m \leqq 0) = P((Y-m)/\sqrt{m} \leqq 0)$ which converges to $\frac{1}{2}$ as $m \to \infty$ by the Central Limit Theorem. Using Stirling's approximation, we see that $e^m m!/m^m \sim \sqrt{2\pi m}$ as $m \to \infty$ from which it follows that $E(X) \sim \sqrt{\pi m/2}$ as $m \to \infty$.

II. *Solution by R. J. Dickson, Lockheed Palo Alto Research Laboratory.* There are m^{k+1} sample sequences of length $k+1$ and $m(m-1) \cdots (m-k+1)k$ of them lead to termination at that length. The probability of termination at length $k+1$ is therefore $P_{k+1} = m!k/(m-k)!m^{k+1}$ for $k = 1, 2, \cdots, m$ and zero for other values of k. Hence the expectation $E(m)$ has the stated form.

Making the substitution

$$\frac{1}{m^{k+1}} = \frac{1}{k!}\int_0^\infty e^{-ms}s^k ds,$$

in the expression for $E(m)$, then interchanging the order of summation and integration, and integrating by parts we have

$$E(m) = m^2 \int_0^\infty e^{-ms}(1+s)^{m-1}s^2 ds.$$

Now, making the substitution $t = s - \log(1+s)$ $(s \geqq 0)$ in the above integral, we get

$$E(m) = m^2 \int_0^\infty e^{-mt}s(t)dt,$$

where $s(t)$ denotes the unique positive root of the equation $t = s - \log(1+s)$. The complete asymptotic expansion of $E(m)$ can now be obtained by the method of Watson which formally consists of expanding $s(t)$ in fractional powers of t and integrating term by term. (Details of this method and its justification can be found in E. T. Copson, *An Introduction to the Theory of Functions of a Complex Variable*, 1935, pp. 218–219.) By using the expansion of $s(t)$ as given by Copson (*ibid.* p. 221) we find that the first three terms of the expansion are

$$E(m) \sim \sqrt{\frac{\pi}{2}}\, m^{1/2} + \frac{2}{3} + \frac{1}{12}\sqrt{\frac{\pi}{2}}\, m^{-1/2} + \cdots.$$

As an example, these three terms yield the approximation 3.22550 for the value of $E(4)$; since $E(4)=3+7/32=3.21875$, this approximation is in error by only 0.21%.

Also solved by Frederick Carty, M. Hirshhorn (Scotland), Eric Langford, Harry Lass & Peter Gottlieb, and an unnamed solver. Partial solutions by R. M. Anderson, M. Becker & R. Garfield, Keith Berry, R. C. Bollinger, Michael Goldberg, Ellen Hertz, P. M. Kannan, G. S. Rogers, F. G. Schmitt, Jr., Kim Scorp, Jim Tattersall, and P. H. Young. P. G. Kirmser and E. F. Schmeichel noted the mistake in the original statement of the problem but did not submit solutions.

The unnamed solver refers to B. Harris' *Probability distributions related to random mappings*, Ann. Math. Stat., 31(1960) 1045–1062. By using the technique of Dickson, we can show that
$E(m) \sim \sqrt{\pi/2}\, m^{1/2} + 2/3 + (1/12)\sqrt{\pi/2}\, m^{-1/2} - (4/135)\, m^{-1} + (1/288)\sqrt{\pi/2}\, m^{-3/2} \pm \cdots$.

Two Fibonacci Congruences

E 2264 [1970, 1008]. *Proposed by Erwin Just, Bronx Community College*

If F_k is the kth Fibonacci number ($F_k=F_{k-1}+F_{k-2}$, $F_1=F_2=1$), prove the following congruence relations:

$$\text{(a)} \qquad F_{7m}(1-F_{7k+1}) \equiv F_{7k}(1-F_{7m+1}) \pmod{377}$$

$$\text{(b)} \qquad 5(F_{k-1}^2+F_{k+1}^2) \equiv 2(-1)^{k+1} \pmod{(F_{k-1}+F_{k+1})}.$$

Solution by D. M. Bloom, Brooklyn College. Since $F_7=13$ divides both F_{7m} and F_{7k}, relation (a) holds mod 13. Hence to prove (a) it suffices to show that

$$\text{(1)} \qquad F_{7m}(1-F_{7k+1}) \equiv F_{7k}(1-F_{7m+1}) \pmod{29}.$$

Since $F_0\equiv F_{14}$ and $F_1\equiv F_{15}$ (all congruences are now mod 29) it follows by induction that $F_n\equiv F_{n+14}$ for all n and hence that $F_n\equiv F_{n+14q}$ for all n and q. If m and k are both odd this implies that $F_{7m}\equiv F_{7k}$ and $F_{7m+1}\equiv F_{7k+1}$; if at least one of m and k is even (say m) we obtain similarly $F_{7m}\equiv F_0=0$ and $F_{7m+1}\equiv F_1=1$. In either case (1) follows, proving (a).

Congruence (b) is a consequence of the following known identity:

$$\text{(2)} \qquad 5(F_{k-1}^2+F_{k+1}^2) = 3(F_{k-1}+F_{k+1})^2+2(-1)^{k+1}.$$

Also solved by Anders Bager (Denmark), L. Carlitz, R. S. Castroll, M. K. Chowdury, George Drazen & the University of Puget Sound Problems Seminar, R. Garfield, M. G. Greening (Australia), Robert Heller, M. Hirschhorn (Scotland), Stephen Hoffman, R. L. Jow, L. Kuipers, Simeon Reich (Israel), Edith V. Sloan, T. E. Stanley (England), The St. Olaf College Students, E. M. Stone, C. C. Yalavigi & R. G. Kulkarni (India), Stuart Yarus, A. Zujus, and the proposer.

Editorial Note. Bloom gives no reference for his identity (2), but it follows easily from the basic identity

$$(*) \qquad F_k^2 - F_{k-1}F_{k+1} = (-1)^{k+1}.$$

Consider $3(F_{k-1}+F_{k+1})^2+2(-1)^{k+1}$. Expanding out the square and substituting in (*) gives

$$3F_{k-1}^2+3F_{k+1}^2+6F_{k-1}F_{k+1}+2F_k^2-2F_{k-1}F_{k+1}.$$

If we now combine terms and apply the fundamental identity for Fibonacci numbers, namely $F_{k+1} - F_{k-1} = F_k$, this becomes

$$3F_{k-1}^2 + 3F_{k+1}^2 + 4F_{k-1}F_{k+1} + 2(F_{k+1} - F_{k-1})^2 = 5(F_{k-1} + F_{k+1})^2,$$

which establishes (2).

More about Magic Star Polygons

E 2265 [1970, 1106]. *Proposed by N. M. Dongre, Sydenham College, India*

Let a regular star polygon be constructed by dividing a circle into n equal parts and by drawing the chords joining alternate points of division. Each of the n chords will carry four points of intersection. It is desired to assign the integers $1, 2, \cdots, 2n$ to the $2n$ points of intersection so as to have a magic star polygon (i.e., the sum of the four numbers on each chord is constant: see Problem E 2092 [1969, 557]). Prove that a necessary condition for the existence of a magic star polygon is that $n > 5$. Is this condition sufficient?

Solution by the proposer. For the magic star pentagon shown in Figure 1, let s denote the sum of the elements along any chord. Then $5s$ is equal to twice the sum of all of the elements, so that $5s = 110$ and hence $s = 22$. By subtracting from the sum of all of the elements (viz. 55) the sum of those elements on the two chords through a, one obtains $a + 11 = b + e + h$. Similarly it can be seen that $f + 11 = d + j + h$, etc.

Consider the placement of 10 and 1. If they lie on the same chord, say at b and e, then $a + 11 = 11 + h$, implying that $a = h$; but we cannot have two equal elements, so that 10 and 1 must lie on different chords, say at a and b. Then $10 + 11 = 1 + e + h$, so that $e + h = 20$, which is impossible under the given conditions. Thus there can exist no magic star pentagon, and so $n > 5$ is a necessary condition.

As regards the sufficiency of the condition, I have been able to construct by trial and error magic star polygons for $n = 6$ (see Figure 2), 7, and 9. I do not know of any general method or if a solution exists for other values of n.

Also solved by M. A. Sharaf (Saudi Arabia).

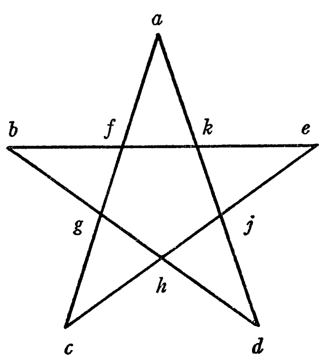


FIG. 1

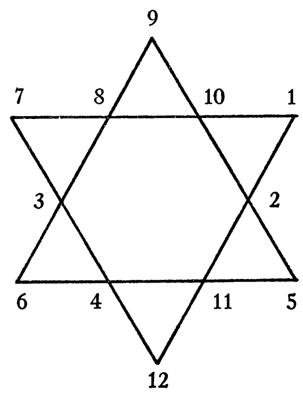


FIG. 2

Two Vector Sums

E 2266 [1970, 1106]. *Proposed by Leon Gerber, St. John's University, Brooklyn, N. Y.*

If the lines joining the vertices of a simplex $\{A_i\}_0^n$ to its centroid M meet the circumsphere again in the points $\{B_i\}_0^n$, then

(a) $$A_iM \cdot MB_i = \sum A_iA_j^2/(n+1)^2, \qquad i, j = 0, \cdots, n; i \neq j.$$

(b) $$\sum A_iM/MB_i = n+1.$$

Solution by Manny Yothers, Lower Stillwater College. The conditions of the problem are much more restrictive than necessary. We shall prove the following more general result: *Let $A_0, A_1, \cdots, A_n$ be $n+1$ points in k-dimensional space, where $k, n \geq 1$. Assume only that the $n+1$ points lie on some k-sphere S, and that they are not all the same. Let M be the barycenter of $A_0, A_1, \cdots, A_n$ and suppose that S has center O and radius r. Let B_i be the intersection of the extension of MA_i and S. Then*

(a) *For all i,*

$$(A_iM)(MB_i) = r^2 - (OM)^2 = \frac{1}{2(n+1)^2} \sum_{i=0}^{n} \sum_{j=0}^{n} (A_iA_j)^2,$$

(b) $$\sum_{i=0}^{n} \frac{A_iM}{MB_i} = n+1.$$

Part (a) deserves some comment. The stated problem is in error when it omits the factor $\frac{1}{2}$ on the right-hand side. Further, it makes no difference in the sum to require that $i \neq j$ since $A_iA_i = 0$.

To prove the general result, assume without loss of generality that O is the origin. Define the vectors $\mathbf{a}_i = OA_i$, $\mathbf{b}_i = OB_i$, and $\mathbf{m} = OM$. Note that $(n+1)\mathbf{m} = \sum \mathbf{a}_i$, and that no $\mathbf{a}_i = \mathbf{m}$. Fix i and consider the two-dimensional circle of radius r formed by the intersection of S and the plane determined by O, A_i, and B_i. We recall from plane geometry that if AB and CD are chords of a circle which intersect at a point M in the circle, then $(AM)(MB) = (CM)(MD)$. It follows then that

$$(A_iM)(MB_i) = (r + OM)(r - OM) = r^2 - (OM)^2.$$

(If $k = 1$, this is trivial.) Now consider the following sum:

$$\sum_{i=0}^{n} \sum_{j=0}^{n} (A_iA_j)^2 = \sum_{i=0}^{n} \sum_{j=0}^{n} (\mathbf{a}_i - \mathbf{a}_j) \cdot (\mathbf{a}_i - \mathbf{a}_j).$$

If we expand out the dot product and use the facts that $\mathbf{a}_i \cdot \mathbf{a}_j = r^2$ and that $\sum \mathbf{a}_i = (n+1)\,\mathbf{m}$, then it is easily seen that

$$\sum_{i=0}^{n} \sum_{j=0}^{n} (A_iA_j)^2 = 2(n+1)^2(r^2 - |\mathbf{m}|^2),$$

which proves (a).

To show (b), consider the following sum:

$$\sum_{i=0}^{n} \frac{A_iM}{MB_i} = \sum_{i=0}^{n} \frac{(A_iM)^2}{(A_iM)(MB_i)}$$

$$= (r^2 - |\boldsymbol{m}|^2)^{-1} \sum_{i=0}^{n} (\boldsymbol{a}_i - \boldsymbol{m})\cdot(\boldsymbol{a}_i - \boldsymbol{m}).$$

Expand out the dot product again and use the same facts; this establishes (b).

Also solved by Anders Bager (Denmark), J. C. Binz (Switzerland), Michael Goldberg, M. G. Greening (Australia), Peter Kornya, and the proposer.

ADVANCED PROBLEMS

All solutions of Advanced Problems should be sent to J. Barlaz, Rutgers—The State University, New Brunswick, N. J. 08903. Solutions of Advanced Problems in this issue should be typed (with double spacing) on separate, signed sheets and should be mailed before February 29, 1972. Contributors (in the United States) who desire acknowledgment of receipt of their solutions are asked to enclose self-addressed, stamped postcards.

An asterisk (*) means neither the proposer nor the editors supplied a solution.

5820. *Proposed by Julio Cano, Findlay College, Findlay, Ohio*

Let K be a compact subset of the line and f a continuous function from K to K. Suppose that $x_0 \in K$ has the property that every cluster point of the sequence $\{f^n(x_0)\}$ is a fixed point of f. Show that the sequence is convergent. Show also that this result fails in two dimensional Euclidean space.

5821. *Submitted by Eric Langford, University of Maine*

Let I denote the unit interval $[0, 1]$. (a) Suppose that E is a Lebesgue-measurable subset of I such that $0 < m(E) < 1$. Show that

$$\text{(i)}\ \inf_J \frac{m(E \cap J)}{m(J)} = 0,$$

$$\text{(ii)}\ \sup_J \frac{m(E \cap J)}{m(J)} = 1,$$

where the supremum and infimum are taken over the class of all nontrivial, proper subintervals J of I.

(b) Does there exist a set E such that for every J, $0 < m(E \cap J) < m(J)$? I.e., does there exist a set E which meets every nontrivial interval in a set of positive measure, and whose complement $I \backslash E$ does likewise?

5822*. *Proposed by Joseph Malkevitch, York College, New York City*

Does there exist a planar simple closed curve K such that every line through every point in the interior of K meets the boundary of K in precisely $2r$ points (r an integer $\geqq 2$)?

5823. *Proposed by J. T. Arnold, Virginia Polytechnic Institute, and J. W. Brewer, University of Kansas*

Let D be an integral domain with identity. Show that if the primary ideals of D are linearly ordered under $\subseteqq$ and if D satisfies the ascending chain condition on prime ideals, then D is a valuation ring.

5824. *Proposed by M. F. Neuts, Purdue University*

Show that for every finite complex number u,

$$\lim_{n\to+\infty} \exp(-u\gamma\sqrt{n})\cdot\Gamma^n\left(1-\frac{u}{\sqrt{n}}\right)=\exp\left(\frac{\pi^2u^2}{12}\right),$$

where $\Gamma(\cdot)$ is the gamma function and γ is Euler's constant.

5825. *Proposed by Erwin Just, Bronx Community College*

Assume that α and β are real numbers, $\beta\neq 0$, such that $\alpha+\beta i$ is a zero of $f(x)$, a cubic polynomial with rational coefficients. If $g(x)$ is the minimal polynomial (with rational coefficients) of βi, can any of the zeros of $g(x)$ be real?

SOLUTIONS OF ADVANCED PROBLEMS

Cycle Decompositions in the Symmetric Group

5751 [1970, 775]. *Proposed by Marvin Marcus, University of California at Santa Barbara*

Let S_m denote the symmetric group of degree m. For $\sigma\in S_m$ let $c(\sigma)$ be the number of cycles (including cycles of length 1) in the distinct cycle decomposition of σ. Prove that for $1\leqq m\leqq n$,

$$\binom{n}{m}=\frac{1}{m!}\sum_{\sigma\in S_m}(\operatorname{sgn}\sigma)n^{c(\sigma)}. \tag{1}$$

Also prove that for any positive integers m and n,

$$\binom{n+m-1}{m}=\frac{1}{m!}\sum_{\sigma\in S_m}n^{c(\sigma)}. \tag{2}$$

Solution by Ralph Freese, California Institute of Technology.

Consider

$$x(x-1)(x-2)\cdots(x-m+1)=\sum_{\sigma\in S_m}(\operatorname{sgn}\sigma)x^{c(\sigma)}, \tag{1'}$$

$$x(x+1)(x+2)\cdots(x+m-1)=\sum_{\sigma\in S_m}x^{c(\sigma)}, \tag{2'}$$

which imply (1) and (2) under the substitution of n for x. Both sides of (1′) and (2′) are polynomials which are to agree at infinitely many integers.

We prove (1′) by induction: $m=1$ is easy. Assume (1′) holds for $m=k$ so that

$$x(x-1)\cdots(x-k+1)=\sum_{\sigma\in S_k}\operatorname{sgn}(\sigma)x^{c(\sigma)}.$$

Hence

$$x(x-1)\cdots(x-k+1)(x-k) = (x-k)\sum_{\sigma\in S_k}\operatorname{sgn}(\sigma)x^{c(\sigma)}$$
$$= \sum_{\sigma\in S_k}\operatorname{sgn}(\sigma)x^{c(\sigma)+1} - k\sum_{\sigma\in S_k}\operatorname{sgn}(\sigma)x^{c(\sigma)}.$$

Now extend the σ's in the first of the two sums to S_{k+1} by defining $\sigma(k+1)=k+1$. Choose r, $1\leqq r\leqq k$, and extend the σ's of the second sum by multiplying by the transposition $(r, k+1)$, so that $\sigma(k+1)=r$. Since there are k sums we do this for each r such that $1\leqq r\leqq k$. Note that in the first case the extended σ has one more cycle than the original σ, but the same sign as σ; in the second case the number of cycles remains unchanged but the sign changes. Denoting the extended σ by σ', we have for the right hand side of the above equation:

$$\sum_{\sigma\in S_k}\operatorname{sgn}(\sigma')x^{c(\sigma')} - k\sum_{\sigma\in S_k}(-\operatorname{sgn}(\sigma'))x^{c(\sigma')}$$
$$= \sum_{\substack{\rho\in S_{k+1}\\ \rho(k+1)=k+1}}\operatorname{sgn}(\rho)x^{c(\rho)} - \sum_{r=1}^{m}\sum_{\substack{\rho\in S_{k+1}\\ \rho(k+1)=r}}(-\operatorname{sgn}(\rho))x^{c(\rho)} = \sum_{\rho\in S_{k+1}}\operatorname{sgn}(\rho)x^{c(\rho)}$$

as was to be shown. (A similar induction proves (2).)

It can also be shown that (1) and (2) are generalizable to arbitrary irreducible characters of S_m as follows: let $p_1\geqq p_2\geqq\cdots\geqq p_r\geqq 1$ be a partition of m. That is, $p_1+\cdots+p_r=m$. Let χ be the character corresponding to this partition (via the Young diagram). Then the generalized formula is

$$\chi(e)^2 p_1!p_2!\cdots p_r!\binom{n+p_1-1}{p_1}\binom{n+p_2-2}{p_2}\cdots\binom{n+p_r-r}{p_r}$$
$$= \frac{\chi(e)}{m!}\sum_{\sigma\in S_m}\chi(\sigma)n^{c(\sigma)}.$$

Also solved by D. M. Bloom, L. Carlitz, L. E. Clarke (England), D. M. Cohen, Red Cougar, R. L. Davis, D. Ž. Djoković, Neal Felsinger, M. G. Greening (Australia), A. A. Jagers (Holland), Jürg Rätz (Switzerland), Simeon Reich (Israel), David Roselle, M. F. Smiley, Richard Stanley, R. K. Tamaki, and the proposer.

Editorial Notes. (1) Carlitz, Davis, Roselle, and Jagers point out that the result follows easily from the fact that the number of permutations of S_m with k cycles is equal to $s(m,k)$, the Stirling number of the first kind, defined by $x(x-1)\cdots(x-m+1)=\sum_{k=1}^{m}s(m,k)x^k$. (See Riordan, *Combinatorial Analysis*, Ch. 4.)

(2) The solution by Rätz follows directly from formulas in his paper, *Zur Zerlegung von Permutationen in elementfremde Zyklen*, Elemente d. Math., 22 (1967), p. 13 ff.

Modified Euler Product

5756 [1970, 891]. *Proposed by J. H. Westbrooke, Evanston, Ill.*

Let $q_1=13$, $q_2=17$, and $q_3, q_4,\cdots$ be the successive primes of the form $4k+1$. Find the least upper bound of the sequence whose nth term is

$$1-\prod_{k=1}^{n}\left(1-\frac{2}{q_k}\right).$$

Solution by J. R. Smart, University of Wisconsin. Since the series $\sum_{p\equiv 1(4)} 1/p$ diverges, the infinite product $\prod_{p\equiv 1(4)} (1-2/p)$ diverges to 0. Thus

$$\lim_{n\to\infty}\left[1-\prod_{k=1}^{n}\left(1-\frac{2}{q_k}\right)\right]=1.$$

Also solved by W. C. Waterhouse, and by the proposer.

Editorial Note. Waterhouse notes that 2 may be replaced by any positive integer, and $\{q_k\}$ by the primes in any arithmetic progression—or more generally, by any set for which $\sum 1/q_k=\infty$.

An Integral-Differential Equation

5757 [1970, 891]. *Proposed by R. A. Struble, North Carolina State University*

Does there exist a solution of the boundary value problem:

$$\frac{dy}{dt}+\sin y=\int_{t/2}^{t}[1+y^2(s)]\sin s\,ds,\qquad y(0)=y(1)?$$

If so, how many?

Solution by the proposer. The answer is yes; indeed there are at least two such solutions. Let C be the complete metric space of continuous mappings from the closed interval $I=[0, 1]$ into itself with the sup metric d. Then for each $y_0\in I$ and $x\in C$, there is a unique solution in C of the initial value problem

$$\text{(DE)}\qquad \frac{dy}{dt}+\sin y=\int_{t/2}^{t}[1+x^2(s)]\sin s\,ds,\qquad y(0)=y_0,$$

(which depends continuously upon the initial value y_0). To see this, we note that the differential equation in (DE) is globally Lipschitzian and that any solution with initial value $y_0\in I$ is confined by the upper and lower boundaries so that when $y=1$, $dy/dt\leqq -\sin 1+2\int_{t/2}^{t}\sin s\,ds<0$; and when $y=0$, $dy/dt\geqq\int_{t/2}^{t}\sin s\,ds>0$, for $0<t\leqq 1$. Now let $y_0\in I$ be fixed. For each $x\in C$, let $y=T(x)$ be the unique solution of (DE). Then $T(C)\subseteq C$. We shall show that T is a contraction on C and thus has a unique fixed point. To this end let $y_1=T(x_1)$, $y_2=T(x_2)$ and $\Delta y=|y_1-y_2|$, where $x_1, x_2\in C$. Then we have

$$y_1(t)-y_2(t)=\int_0^t[\sin y_1(s)-\sin y_2(s)]ds+\int_0^t\int_{\tau/2}^{\tau}[x_1^2(s)-x_2^2(s)]\sin s\,ds\,d\tau,$$

and from this we readily obtain the inequalities

$$\begin{aligned}0\leqq \Delta y(t)&\leqq\int_0^t\Delta y(s)ds+2d(x_1,x_2)[2\sin(t/2)-\sin t]\\&\leqq\int_0^t\Delta y(s)ds+d(x_1,x_2)/4,\qquad\text{for } 0\leqq t\leqq 1.\end{aligned}$$

Thus, by Gronwall's lemma, $\Delta y(t) \leqq d(x_1, x_2)e^t/4$, for $0 \leqq t \leqq 1$, and so $d(y_1, y_2) \leqq \frac{1}{4}ed(x_1, x_2)$. The unique fixed point in C of T is a solution of the original integral-differential equation with initial value y_0.

If in the above argument we let $y_1 = x_1$ and $y_2 = x_2$ be the unique fixed points respectively for initial values $y_{01}, y_{02} \in I$, then we readily obtain the inequalities

$$0 \leqq \Delta y(t) \leqq \int_0^t \Delta y(s)ds + d(y_1, y_2)/4 + |y_{01} - y_{02}|, \qquad \text{for } 0 \leqq t \leqq 1,$$

and so, $d(y_1, y_2) \leqq [d(y_1, y_2)/4 + |y_{01} - y_{02}|]e$, or what is the same, $d(y_1, y_2) \leqq 4e|y_{01} - y_{02}|/(4-e)$. Thus the solutions in C of the original integral-differential equation depend continuously upon their initial values. In particular, their terminal values, which all lie in I, depend continuously upon their initial values and so there exists a solution in C of the given boundary value problem. By considering solutions of (DE) with initial values $y_0 > 1$, we can show that there is necessarily another solution of the boundary value problem.

Best Radii of Univalency

5758 [1970, 891]. *Proposed by W. O. Egerland, Research and Development Center, Aberdeen Proving Ground, Md.*

A theorem of E. Landau states: If $f(z)$ is analytic and bounded by 1 in the unit disk, $|z| < 1$, and satisfies the conditions $f(0) = 0$ and $f'(0) = a$, $0 < a < 1$, then $f(z)$ is univalent in the disk $|z| < \rho = a^{-1}(1 - (1-a^2)^{1/2})$. What is the corresponding ρ if, in addition, $f(z) \neq 0$ for $z \neq 0$?

Solution by M. R. Ziegler, University of Kentucky. The function $f(z)/z$ is analytic, nonzero, and bounded by 1 for $|z| < 1$; hence

$$p(z) = \frac{\log f(z)/z}{\log a}$$

is analytic in $|z| < 1$ and satisfies the conditions $\operatorname{Re}[p(z)] > 0$, $p(0) = 1$. It is well known that for such functions, $|p'(z)| \leqq 2/(1-r)^2$. Thus, letting $d = \log a$, $-\infty < d < 0$, we can write

$$\begin{aligned} \operatorname{Re}\left[\frac{zf'(z)}{f(z)}\right] &= 1 + d \operatorname{Re}[zp'(z)] \\ &\geqq 1 + d|zp'(z)| \geqq \frac{1 - 2(1-d)r + r^2}{(1-r)^2}. \end{aligned}$$

Then $f(z)$ will map $|z| < r$ onto a starlike univalent domain whenever $\operatorname{Re}[zf'(z)/f(z)] > 0$ or, equivalently, whenever

$$|z| < \rho = 1 - d - \sqrt{d^2 - 2d}.$$

Furthermore, the function $g_0(z) = z\exp[d(1+z)/(1-z)]$ satisfies the hypothesis,

and a calculation shows $g_0'(\rho)=0$; hence $g_0(z)$ is not univalent in any disk $|z|<r$ if $r>\rho$.

Also solved by the proposer.

Regular Rings

5759 [1970, 1015]. *Proposed by Robert Raphael, McGill University*

Let S be a commutative ring with unity, and let R be a subring of S containing the unity of S. Assume furthermore that S has no nilpotent elements (other than 0), and that each element of S satisfies a monic polynomial equation with coefficients from R. Recall that a ring is regular (in the sense of von Neumann) if for each x there is a y such that $x=x^2y$. Show that if R is regular then S is regular as well.

Solution by A. A. Jagers, Technische Hogeschool Twente, Enschede, Holland. We shall prove by induction on k that all elements of S which satisfy a monic polynomial equation of degree k with coefficients in R and (for technical reasons) with constant term equal to 0, are regular. The case where $k=1$ is trivial. Suppose the regularity has been proved for all $k<n$ and let $x\in S$ be such that $x^n+a_{n-1}x^{n-1}+\cdots+a_1x=0$ with $a_i\in R(1\leqq i\leqq n-1)$. Let b_1 satisfy $(-a_1)^2b_1=-a_1$. Put $i=-a_1b_1$ and $j=1-i$. Then both i and j are idempotents and by appropriate multiplication one has $(ix)^2b_1f(ix)=ix$ and $xg(jx)=0$, where $f(X)=X^{n-2}+a_{n-1}X^{n-3}+\cdots+a_2$ and $g(X)=Xf(X)$. This shows on the one hand that ix is regular, and on the other hand, because $X|g(X)$, that $g^2(jx)=0$. Since the only nilpotent element in S is 0 it follows that $g(jx)=0$, and thus, since the degree of g is less than n, that jx is regular as well. Hence x is regular.

Also solved by W. D. Blair, J. W. Brewer & E. A. Rutter, D. J. Fieldhouse, C. W. Kohls, Surjeet Singh (India), K. C. Smith, R. B. Tarsy, E. T. Wong, W. C. Waterhouse, and the proposer.

Editorial Note. Fieldhouse, Tarsy, Waterhouse, and Wong, each note that the result follows directly from the following two theorems to be found in Bourbaki, *Algèbre Commutative*, Atiyah-MacDonald, *Commutative Algebra*, and I. Kaplansky, *Commutative Rings:* (A) R is regular if and only if (1) R has no nonzero nilpotents, (2) each prime ideal of R is maximal; and (B) Since R is integral over S, R has all its prime ideals maximal if and only if S does.

Uncountable Partition of [0, 1]

5760 [1970, 1015]. *Proposed by D. K. Kosaka, Dallas, Texas*

Without using the axiom of choice, or equivalent, decompose the unit interval [0, 1] into an uncountable class of disjoint sets, each of which is uncountable and everywhere dense in the unit interval [0, 1].

I. *Solution by J. R. Isbell, State University of New York at Buffalo.* Expand the numbers x in [0, 1] as non-terminating proper decimals $.x_1x_2\cdots$ in some definite way, and define $(x, y)\in R$ if $x_i=y_i$ for all composite i except finitely many. Then R is an equivalence relation; so its equivalence classes partition [0, 1]. Each is dense since it has elements with any specified first n digits. Each is uncountable since it has elements with arbitrary digits x_p for prime p.

If the set S of equivalence classes were countable, so would be the domain of any injection into S; so it suffices to exhibit an uncountable subset T of $[0, 1]$ no two of whose elements are R-related. Let T be the set of all decimals x all of whose digits are 1 or 2 and for which the i^j-th digit repeats the i-th digit and $x_1 = 1$. Then T is uncountable since the i-th digits for i not 1 and not a power are arbitrary in $\{1, 2\}$; no two distinct elements are R-related since any difference repeats at infinitely many composite places.

II. *Solution by Richard Stanley, Massachusetts Institute of Technology.* If $x \in [0, 1]$, let $f_x(n)$ be the number of 1's appearing in the first n digits of the binary expansion of x (where we choose, say, the expansion with infinitely many 0's). Let $g(x) = \lim\inf_{n\to\infty} f_x(n)/n$, and for each $\alpha \in [0, 1]$, define $S_\alpha = \{x \mid g(x) = \alpha\}$. Clearly the S_α's decompose $[0, 1]$ into an uncountable class of disjoint sets. Moreover, each S_α is dense in $[0, 1]$ since for any n, any alteration of the first n digits of the binary expansion of x does not affect the value of $g(x)$. Finally, each S_α is uncountable, since for any sequence $n_1 > n_2 > \cdots$ of positive integers satisfying $n_i/i \to \infty$, we can alter arbitrarily the n_i-th digits of the binary expansion of x, for each $i = 1, 2, \cdots$, without affecting the value of $g(x)$.

Moreover, since almost all $x \in [0, 1]$ satisfy $\lim_{n\to\infty} f_x(n)/n = \frac{1}{2}$, it follows that the Lebesgue measure of each S_α is 0, except $\mu(S_{1/2}) = 1$.

Also solved by M. A. Archer, R. C. Brown, Burgess Davis, Giuseppe DeMarco, Steve Ferry, G. J. Foschini, J. W. Grossman, Jan Hejcman (Czechoslovakia), Dennis Henkel, R. B. Israel, C. G. Jockusch, S. A. Kalilow, L. E. Mattics, J. C. Morgan II, Pearl Olson, Nicholas Passell, Roger Purves, Donald Quiring, D. E. Sanderson, Moshe Shimrat, and the proposer.

Morgan refers to a decomposition given by P. Mahlo in Schoenflies, *Entwicklung der Mengenlehre und ihrer Anwendungen*, p. 343, ff. Foschini refers to his paper, *The Pathology of functions which are almost everywhere constant*, Bull. Math. de la Soc. Sci. Math. de la R. S. Roumanie, VII (59), nr. 3, (1967) p. 39, where two proofs appear, the first by Knaster and Kuratowski, which is similar to the second solution above.

UNSOLVED PROBLEMS

This list supplements the list printed on p. 711 of the June-July 1969 issue of this MONTHLY. Any comments or solutions will be welcomed by the editors.

5575 [1968, 790]	5606 [1968, 686]	5643 [1968, 1125]
5580 [1968, 414]	5608 [1968, 686]	5670 [1969, 423]
5589 [1968, 415]	5619 [1968, 792]	5671 [1969, 565]
5599 [1968, 553]	5625 [1968, 911]	5687 [1969, 835]
5603 [1968, 685]	5626 [1968, 911]	5702 [1969, 1152]

and E 2133 [1968, 1112].

Problems in the 1969 list for which solutions have since been received are 5415, E 1847, E 1903, E 1905, E 1917, E 1959, E 1980.

REVIEWS

EDITED BY J. ARTHUR SEEBACH, JR. AND LYNN A. STEEN

with the assistance of the mathematics departments of St. Olaf and Carleton Colleges

COLLABORATING EDITOR FOR FILMS: SEYMOUR SCHUSTER, Carleton College

Printed materials for review should be sent to: Book Review Editor, American Mathematical Monthly, St. Olaf College, Northfield, MN 55057. Films and correspondence relating to films should be sent to Seymour Schuster, Carleton College, Northfield, MN 55057.

All unsigned material is written by the editors. A boldface capital **C** *in the margin indicates that a review is based in part on classroom use. Professors willing to write such a review should inform the editor in order to avoid duplication.*

C *Introduction to Combinatorial Mathematics.* By C. L. Liu. McGraw-Hill, New York, 1968. x+393 pp. $13.50. (Telegraphic Review, March 1969.)

Both this text and the course in which we have (separately) used it testify by their existence to the recent resurgence of interest in "combinatorics."New kinds of problems from modern applications have expanded what was once widely felt to be an amorphous collection of tricks, while at the same time new unifying theories from general mathematics are beginning to impose order in large realms of this domain. Today there is a place in the curriculum for an introduction to combinatorics designed for advanced undergraduates and beginning graduates in mathematics, statistics, and related disciplines.

Our one-semester course makes use of just about half of Liu's book, leaving out the four chapters on graph theory, two on programming, and one on designs (all dealt with elsewhere on our campus), and covering, with various emendations and amplifications, all the material of Chapters 1–5 and 10–11. The first five chapters deal with generating functions, recurrences, and inclusion-exclusion methods, and culminate in a thorough, not to say laborious, treatment of Polya's enumeration theory (along the highly perspicuous lines first laid down by De Bruijn). With our fairly mature students we were able to shortcut cumbersome discussions of elementary abstract mathematics, to provide needed underpinning from the theory of formal power series for Liu's descriptions of generating-function and recurrence methods, and to show in the Rota manner how inclusion-exclusion and rook-polynomial formulas, as well as the basic Polya theory, can be derived in terms of generalized Möbius inversion. We dealt only briefly with Chapter 10 (network flows) and followed the book (apart from an easily corrected error in the proof of the basic theorem) for the next chapter, on matching theory.

Somewhat surprisingly, this is a good text. While the meticulous mathematician may cavil at a great deal in its language and manner, especially when Liu expounds elementary mathematics, the book does not show the fault most typical of expositions of combinatorics some years ago. It seemed then that authors were always tempted to place an inordinate share of the burden of comprehension on the reader; the born combinatorialist, seeing connections not perceptible to ordinary mathematicians, but hard pressed to explain them, too often fell back on what has been called the "Arabian Nights method" and, after

a sequence of increasingly desperate but inconclusive arguments, could only wave his hands and say "Lo!" That Liu's writing usually avoids this traditional weakness surely reflects the influence of the theories with which mathematicians today are trying to unify this recalcitrant material. (Liu, who is an electrical engineer, studied combinatorics with that leading unifier, G.-C. Rota.)

It is difficult to mention the flaws we speak of without implying a disdain we do not feel. We could not write a review for mathematicians without mentioning them, but we and our students were fully agreed in the end that the material of this book is so well chosen, so amply illustrated in examples throughout the text, and so well supported and extended by a large number of good and often challenging problems, that (with appropriate classroom monitoring) it had made an excellent text.

R. L. DAVIS and D. G. KELLY, University of North Carolina

C *Elementary Differential Equations with Linear Algebra.* By Albert L. Rabenstein. Academic Press, New York, 1970. ix+441 pp. $10.50. (Telegraphic Review, August/September 1970.)

This is a well written textbook for an introductory course in differential equations with linear algebra. It is not as advanced as the author's earlier book, *Introduction to Ordinary Differential Equations* (Academic Press, 1966).

This book is used here as a textbook for differential equations. A lecture or two on basic matrix theory and a few lectures on Chapter 3 are sufficient to give a minimal background in linear algebra. This is a very flexible book. There are a variety of ways that one can add to this minimal list of topics for linear algebra. In addition, the book can be used to give a fairly comprehensive introduction to linear algebra, as well as to differential equations.

Most of the topics that one would normally find in an introductory book on differential equations are covered. (Laplace transformations are not mentioned.) Theoretical considerations connected with differential equations are carefully introduced in Chapter I. This theoretical development culminates in the statement and proof of the fundamental (existence and uniqueness) theorem for linear differential equations in Chapters 5 and 8. In Chapter 5 the author introduces "just enough" complex variable theory to operate effectively. First-order systems of differential equations are considered in Chapter 6. Such a system is first solved without using matrices and then as an eigenvalue problem.

As for the linear algebra, again most of the basic topics are covered. The concept of a matrix of row echelon form is not used, and the author does not give sufficient attention to the topic of canonical form. This attitude seems to be reflected in the illustration on page 109 for finding the inverse of a matrix by elementary row operations. Essentially, the illustration consists of listing a matrix and its inverse!

I recommend this book very highly for an introductory course in differential equations with (or without) linear algebra. The two subjects are very successfully integrated—this is not "token integration"!

D. H. TRAHAN, Naval Postgraduate School

Introduction to Computers and Programming. By J. Hellwig. Columbia University Press, New York 1969. 215 pp. $6.60. (Telegraphic Review, August/September 1970.)

This book has been written to introduce the novice to the basic characteristics of computer systems and programming, principally FORTRAN. It is designed for self-study: after reading about the essential facts of each subject, the student is taken step by step through an illustrative example. The main subjects covered in this book are formulation of problems and the use of flow diagrams, storage and number systems, input and output of information, operating systems and system diagnostics.

This is a good, readable introductory text for persons unfamiliar with computers who are interested in a cursory knowledge in the subjects listed above. It should be emphasized that there is very little in this book about programming techniques themselves. My major criticism is that the book lacks a bibliography, especially since further reading is necessary for anyone having more than a dilettant's interest in these subjects.

LEON LEVINE, University of California, Los Angeles

C *Celestial Mechanics.* By Shlomo Sternberg. Benjamin, New York, 1969. 2 vols. 179 pp., 321 pp. $15, $6.95 (P). (Telegraphic Review, April 1970.)

These two volumes were notes for a course given at Harvard during the spring of 1968. "The main purpose of the course was to study recent developments in the subject, principally the ideas inaugurated by Kolmogoroff and developed by Arnold and Moser." (The K.A.M. Theory.) It must be emphasized that these are lecture notes and not a textbook. They reflect the personal taste of the author to such an extent that a more accurate title would be "Topics in Analysis arising from Celestial Mechanics."

The first volume contains a brief but fascinating sketch of some problems in ancient astronomy including a discussion of Ptolemy's theory of epicycles. The mathematical description of epicycles requires quasi-periodic functions. Therefore, a brief survey of the theory of almost periodic functions is appropriate in a book on celestial mechanics. Professor Sternberg, however, develops a short course on the subject which includes the Birkhoff ergodic theorem, the Peter-Weyl theorem and the Bohr approximation theorem. The reviewer used this section in a seminar and found that the proofs in the standard textbooks were far easier to present than those given in Sternberg's notes.

The next topic in Part I is an interesting, but not very clear, introduction to some problems in celestial mechanics including the three body problem, Hill's lunar theory, and even general relativity. Once again, it is important to remember that these are lecture notes, not a textbook. The reader who has difficulty following the material should read *Mathematical Introduction to Celestial Mechanics* by H. Pollard, or *Theory of Orbits* by V. Szebehely and return. Sternberg's point of view is important and deserves careful study.

The second volume is primarily a course in nonlinear functional analysis which requires considerable mathematical maturity and perseverance to under-

stand. The study of flows generated by a smooth vector field on the two dimensional torus leads naturally to the K.A.M. theory. The Nash embedding theorem for Riemannian manifolds "turns out to be an easy consequence of a special case" of this theory, and a proof of it is presented in an appendix. A sketch of the proof of the Birkhoff fixed point theorem and the billiard ball problem are included in the functional analysis chapter.

The concluding chapter treats the three body problem in more detail than the earlier treatment in part one and shows how the K.A.M. theory can be used to resolve some of the deepest problems in celestial mechanics.

In the introduction, Sternberg writes "When these notes were about two-thirds completed, I came into possession of the excellent book *Problèmes ergodiques de la mécanique classique* by V. I. Arnold and A. Avez . . . a delightful introduction to the current ideas in mechanics which very successfully conveys the flavor of the subject. I strongly recommend it to one and all." The reviewer agrees. Nevertheless, *Celestial Mechanics* contains an enormous amount of significant and challenging mathematics, and it is well worth the effort of studying the two volumes.

W. T. KYNER, University of New Mexico

Foundations of Probability. By Alfred Rényi. Holden-Day, San Francisco, 1970. 382 p. $20. (Telegraphic Review, December 1970.)

This is an excellent introductory textbook of probability theory on a graduate level, assuming knowledge of measure theory.

In every detail the book bears the original imprint of the author. Thus, conditional probability spaces, due to the author, are introduced in Chapter 2 ("Probability") immediately after the first chapter on "Experiments." Effective constructions of probability spaces and the combinatorial approach are emphasized, and "information" is treated as a basic notion of the theory. The central Chapter 3 on "Independence" considers also "qualitative" independence. Chapter 4 on "Laws of Chance" is devoted to limit theorems for sequences of independent random variables, while the last Chapter 5 on "Dependence" treats martingales, Markov chains, stable and mixing sequences of events, and exchangeable events.

Each chapter is followed by ten exercises and ten problems which form an integral part of the book and frequently are results of the author's research. Propositions and proofs within the text proper are also frequently due to the author.

As the author points out, the limit theorems are not presented "in their most general form; on the contrary we present these theorems usually in a generality which is just enough to give a clear idea of the importance of the theorem in question."

This work is more than a textbook; it is the late Rényi himself speaking.

M. LOÈVE, University of California, Berkeley

TELEGRAPHIC REVIEWS

Telegraphic reviews are intended to give prompt notice of books, with sufficient information to assist in deciding whether to order an examination copy or suggest library purchase. Possible uses are indicated as follows:

B = college bookstore stock L = library purchase
P = professional reading S = supplementary reading
T = textbook E = teacher education
13 to 18 = freshman to second year graduate level usage
1 to 4 = approximate time in semesters to cover text
* = positive emphasis ? = negative emphasis

Books on high school material (pre-calculus) are denoted REMEDIAL, and normally receive telegraphic reviews only if they are written for college students. Publishers are denoted by the standard abbreviations used in *Books in Print*, which gives complete addresses.

ALGEBRA, P, L, *Lecture Notes in Mathematics-173: Structure of Arbitrary Purely Inseparable Extension Fields*. John N. Mordeson and Bernard Vinograde. Springer-Verlag, 1970, 138 pp, $4.10 (P). Largely a presentation of the infinite degree theory, especially for the case without exponent. Includes a bibliography and reference notes following each chapter. L.C.L.

ALGEBRA, T(13: 2), S, *The Art of Algebra*. B. Abrahamson and M.C. Gray. Rigby Limited, 1971, 592 pp, $6.50. Over two-thirds of the content of this inexpensive volume is commonly covered in high school and beginning calculus courses: elementary set theory, inequalities, mathematical induction, real and complex numbers, series and sequences, conics, polynomials (requires calculus), elementary linear algebra. The remaining chapters on groups and fields are very elementary; rings, integral domains and abstract vector spaces are not mentioned. It includes historical remarks and interludes on the nature of mathematics. L.C.L.

ALGEBRA AND NUMBER THEORY, T(13-14: 1), S, L*, *Linear Algebra with Applications*. Hugh G. Campbell. Appleton-Century-Crofts, 1971, xiii + 375 pp, $10.95. This book should appeal to freshman and sophomore students--particularly non-mathematics majors. The theory is carefully presented at the correct level and the applications, given in subsections at the end of most sections, encompass a broad spectrum of disciplines. Many of these are documented and serve primarily as enticements for further investigations. Thus the bibliography could become an integral part of the course through various independent study projects. Such a format for more advanced linear algebra and modern algebra textbooks would be a valuable addition to the literature. L.C.L.

ALGEBRA AND NUMBER THEORY, T*(13-14: 2), S, L, *First Course in Algebra and Number Theory*. Edwin Weiss. Acad Pr, 1971, xi + 547 pp, $12.95. This respected mathematician argues that the best modern algebra and linear algebra texts have been written at the junior-senior level. This book is written at the freshman-sophomore level, and could be taken simultaneously or even in place of first year calculus. The emphasis is on algebraic questions arising out of number theory. There are four slow-moving, wordy, but rather fully

developed chapters on elementary number theory, rings and domains, congruences and polynomials, and groups. L.C.L.

ALGEBRA AND NUMBER THEORY, P, L. *Lecture Notes in Mathematics-182: Cyclic Difference Sets*. Leonard D. Baumert. Springer-Verlag, 1971, vi + 166 pp, $4.60 (P). A survey of the general theory of these sets which also presents some of the outstanding problems. After providing an introduction, the author takes up existence questions, multipliers and constructuve existence tests, sets of special type (e.g., planar, Hadamard), and families of sets (Singer, Nth power residue). Included also is a table of all 85 known cyclic difference sets with at most 100 members. D.F.A.

ALGEBRA, GEOMETRY, T(16-17: 1), S, P, L. *Finite Reflection Groups*. C.T. Benson and L.C. Grove. Bogden & Quigley, 1971, viii + 110 pp. This book attempts to provide a discussion of finite reflection groups which is as elementary as possible. Though the only prerequisites are a good understanding of linear algebra and a sprinkling of topics from group theory, the presentation remains too condensed and sophisticated for most undergraduates to digest independently. A volume of twice this size would have been more realistic. L.C.L.

ALGEBRA, LINEAR ALGEBRA, T*(14: 1). *Linear Algebra and Analytic Geometry*. Heinrich W. Brinkmann and Eugene A. Klotz. A-W, 1971, xvi + 535 pp, $11.50. This book is designed to follow a first year course in Calculus and is adaptable to courses of varying length. Discussions are informal and entertaining which may be misleading. It will require some mathematical maturity. Exercises are mainly proofs which should develop a facility for writing proofs. The contents deal with abstract vector spaces over the field of real numbers with supplementary "honors projects" extending the material to vector spaces over the complex numbers. The analytic geometry is innovative and it also contains the standard facts about euclidean spaces and quadratic forms. L.L.K.

ALGEBRA AND NUMBER THEORY, P, L. *Lecture Notes in Mathematics-189: Dirichlet Series and Automorphic Forms*. André Weil. Springer-Verlag, 1971, iv + 164 pp, $4.60 (P). The relation between modular forms and Dirichlet series with functional equations was discovered by Hecke. The main result of these notes generalizes a theorem of Weil on modular forms (Math Ann 168, 1967, pp 149-156) which is an extension of work by Hecke. The adele point of view is systematically adopted. R.B.K.

ANALYSIS, FUNCTIONAL ANALYSIS, P, L. *Multipliers, Positive Functionals, Positive-Definite Functions, and Fourier-Stieltjes Transforms. Memoirs of the American Mathematical Society, Number 111*. Kelly McKennon. AMS, 1971, vi + 67 pp, $1.90 (P). An extension of the author's Ph.D. thesis under Edwin Hewitt, this memoir examines relationships between positive linear functionals and multipliers (centralizers) on an algebra without identity; the results obtained are used to develop generalizations of known convergence theorems for positive functionals and are applied to group algebras and to Fourier and Fourier-Stieltjes algebras. D.F.A.

ANALYSIS, TOPOLOGY, P, L. *Lecture Notes in Mathematics-175: On*

Topologies and Boundaries in Potential Theory. Marcel Brelot. Springer-Verlag, 1971, vi + 176 pp, $5 (P). The first of the two parts of this volume develops the general notion of topological thinness and the corresponding fine topology associated with a given space and a cone of positive functions, and gives applications to classical potential theory; extensions to axiomatic theories of harmonic functions are indicated briefly. The second part studies abstract minimal thinness and classical boundary theories (including that of the Martin boundary). Some exercises, many references. D.F.A.

APPLIED MATHEMATICS AND GAME THEORY, T(16-17: 1, 2), S, P, L. *Some Topics in Two-Person Games: Modern Analytic and Computational Methods in Science and Mathematics, Number 22.* T. Parthasarathy and T.E.S. Raghavan. Ed: Richard Bellman. Am Elsevier, 1971, xii + 259 pp, $18. A rather comprehensive treatment of two-person game theory along with some topics from n-person games. Includes generalized minimax theorems, differential games, stochastic games, and a brief treatment of the necessary topology and convexity theory. F.L.W.

APPLIED MATHEMATICS, BIOMATHEMATICS, P*, L*. *Mathematical Topics in Population Genetics.* Ed: Ken-ichi Kojima. Springer-Verlag, 1970, vii + 400 pp, $19.60. This is Volume 1 of a new series of monographs and reports in the area of biomathematics. It contains 13 papers by contributors with strong biological backgrounds and is written from the geneticist's rather than the mathematician's point of view. R.S.K.

APPLIED MATHEMATICS, PHYSICS, P, L. *Lectures on Elementary Particles and Quantum Field Theory. 1970 Brandeis University Summer Institute in Theoretical Physics, Volume 2.* Ed: Stanley Deser, Marc Grisaru, and Hugh Pendleton. MIT Pr, 1970, vii + 502 pp, $6.95 (P). Lectures by Rudolph Haag on observables and fields, by Maurice Jacob on Regge models and duality, by Henry Primkoff on weak interactions, by Michael Reed on the GNS-construction, and by Bruno Zumino on effective Lagrangians and broken symmetries. L.C.L.

APPLIED MATHEMATICS, SPECIAL FUNCTIONS, T(16-17: 1), L. *The Functions of Mathematical Physics, Volume XXIII.* Harry Hochstadt. Wiley, 1971, xi + 322 pp, $17.50. A clear exposition of the following topics: Orthogonal polynomials, the classical orthogonal polynomials, the gamma function, hypergeometric functions, the Legendre Functions, spherical harmonics in p dimensions, confluent hypergeometric functions, Bessel functions, Hill's equation. Applications to certain algebraic equations and the conformal mapping of curvilinear triangles as well as to mathematical physics. R.B.K.

APPROXIMATION THEORY, P, L. *Lecture Notes in Mathematics-187: Topics in Approximation Theory.* Harold S. Shapiro. Springer-Verlag, 1971, viii + 275 pp, $6.40 (P). Notes based on a course given at the Royal Institute of Technology in Stockholm, Fall 1969. Assumes a good background in analysis, but none in approximation theory. Would be useful for a seminar. A good deal of "philosophy" on the nature of approximation theory is included, especially in a chapter on "general aspects of 'degree of approximation'." Good bibliography, frequently referenced by text. R.B.K.

CATEGORY THEORY AND FUNCTIONAL ANALYSIS, P, L. *Compactly Covered Reflections, Extension of Uniform Dualities and Generalized Almost Periodicity: Memoirs of the American Mathematical Society, Number 105.* Michael H. Powell. AMS, 1970, vii + 235 pp, $2.70 (P). Category theory and topological vector spaces combined with applications to analysis. A very thorough introductory chapter, a massive terminology index, and a good symbol index may do much to improve accessibility. J.A.S.

COMBINATORICS, T(18), P*, L*. *Transversal Theory: An Account of Some Aspects of Combinatorial Mathematics.* L. Mirsky. Acad Pr, 1971, ix + 255 pp, $13. Provides a codification of this new area of mathematics, in particular showing its origin in the work of Philip Hall and relating it to the theory of abstract independence. Of special interest is the concluding chapter which indicates current trends in the field, identifies the main areas of research and lists 50 open questions. Exercises and an extensive bibliography are included. R.S.K.

COMBINATORICS, GENERAL, P, L. *Studies in Pure Mathematics: Papers in Combinatorial Theory, Analysis, Geometry, Algebra, and the Theory of Numbers.* Ed: L. Mirsky. Acad Pr, 1971, viii + 276 pp, £5. A tribute to Richard Rado on the occasion of his sixty-fifth birthday. These twenty-seven contributions, by a world-wide group of distinguished mathematicians, reflect the broad range of Rado's work, with special emphasis on combinatorial questions. L.C.L.

COMBINATORICS, GRAPH THEORY, P, L. *Lecture Notes in Mathematics-186: Recent Trends in Graph Theory.* Ed: M. Capobianco, J.B. Frechen and M. Krolik. Springer-Verlag, 1971, vi + 219 pp, $5.30 (P). Proceedings of the graph theory conference held at St. John's University, N.Y., in June, 1970. Contains 26 papers by 31 authors on a wide variety of topics in graph theory. D.F.A.

COMPLEX ANALYSIS, SHEAF THEORY, P, L. *Lecture Notes in Mathematics-172: Gap-Sheaves and Extension of Coherent Analytic Subsheaves.* Yum-Tong Siu and Günther Trautmann. Springer-Verlag, 1971, iv + 172 pp, $4.60 (P). Exposition of gap-sheaves and coherent subsheaf extension following the recent work of W. Thimm, the authors, and others. J.A.S.

DIFFERENTIAL AND INTEGRAL EQUATIONS, T(18: 1, 2), P, L. *Applied Mathematical Sciences, Volume 3: Functional Differential Equations.* J. Hale. Springer-Verlag, 1971, viii + 238 pp, $6.50 (P). Notes from a series of lectures given at UCLA in 1968-69, suitable for an introduction to the subject via a course or through individual study. Equations studied are of the retarded type, and topics presented include existence, stability, and periodicity results. Several examples of special type are studied in detail. No exercises. D.F.A.

DIFFERENTIAL AND INTEGRAL EQUATIONS, P, L. *Asymptotic Behavior of Solutions and Adjunction Fields for Nonlinear First Order Differential Equations. Memoirs of the American Mathematical Society, Number 109.* Walter Strodt and Robert K. Wright. AMS, 1971, 284 pp, $3 (P). The authors consider nonlinear algebraic differential equations where the polynomials involved have coefficients which are analytic in a sector of the complex plane, are asymptotically

equivalent to logarithmic monomials, and are included in a differential field in which every element has prescribed asymptotic behavior; solutions lying in an asymptotically well-behaved extension of this field are then sought. The main result concerns the general first-order case. D.F.A.

DIFFERENTIAL AND INTEGRAL EQUATIONS, S, P, L. *Lectures in Applied Mathematics, Volume III: Partial Differential Equations.* Lipman Bers, Fritz John and Martin Schechter. AMS, 1964, xiii + 343 pp, $12.50. The third printing of this portion of the proceedings of a 1957 summer seminar on applied mathematics at the University of Colorado. Chapters on hyperbolic and parabolic equations by John and on elliptic equations by Bers and Schechter, with a brief supplement on eigenvalue expansions by Lars Gårding and one on parabolic equations by the late A.N. Milgram. Sole prerequisites: a standard course in real and complex analysis and familiarity with introductory functional analysis. D.F.A.

DIFFERENTIAL AND INTEGRAL EQUATIONS, P, L. *Lecture Notes in Mathematics-183: Analytic Theory of Differential Equations.* Ed: P.F. Hsieh and A.W.J. Stoddart. Springer-Verlag, 1971, vi + 225 pp, $5.80 (P). Proceedings of a conference held at Western Michigan University in the spring of 1970 which concerned analytic theory and other topics in ordinary and partial differential equations. Contains a long paper by M. Iwano on bounded solutions and stable domains of nonlinear ordinary differential equations. Other papers are by Bank, Benzinger, Cesari, Coddington, Colton and Gilbert, Comstock, Gollwitzer, Hale, Harris, Hoppensteadt, Kazarinoff, Narayan and Stengle, Sánchez, Schubert, Sibuya, Turrittin, and Wasow. D.F.A.

EDUCATION, GENERAL, T, E. *Modern Syllabus Algebra.* D.G.H.B. Lloyd. Pergamon Pr, 1971, v + 240 pp, $2.25 (P). Designed to relate "traditional" mathematics with the "modern", particularly for teachers trained before the changes in outlook in school mathematics (about 1960). Exercises with answers. L.C.L.

FOUNDATIONS, P, L. *Studies in Logic and The Foundations of Mathematics, Volume 61: Logic Colloquium '69, Proceedings of the Summer School and Colloquium in Mathematical Logic, Manchester, August 1969.* Ed: R.O. Gandy and C.M.E. Yates. North-Holland, 1971, xiv + 451 pp, $15.60. Most of the 21 papers in this volume are concerned with set theory and recursion theory. Six of the papers are on generalizations of recursion theory and include a 60 page survey of this field and its applications by Kreisel. L.C.L.

FOUNDATIONS, P, L. *Grundlagen der Mathematik II.* D. Hilbert and P. Bernays. Springer-Verlag, 1970, xiv + 561 pp, $24.30. This second edition differs from the first mainly in the addition of some consistency proofs of the formalism of number theory of Kalmar and Ackermann. Besides this, several proofs and discussions of proofs (the preface to the second edition mentions 7 such instances) are either simplified or augmented so as to provide better clarification. The subject index has been expanded and an author index added. L.C.L.

*GENERAL, T**(15-16: 2). *Set Theory and Topology.* Philip Nanzetta

and George E. Strecker. Bogden & Quigley, 1971, ix + 117 pp. Lots of good topology in a few pages, partly because there are only three proofs total in this Moore-type text. Extensive material on axiomatic set theory leads through a brief motivational section on the Real line to a solid section on general topology. Good examples, well organized theory, and literate authors make for a very nice "do-it-yourself" topology course. J.A.S.

GENERAL, T(13: 1), *Mathematics for a Liberal Education.* Merlin M. Ohmer. A-W, 1971, vi + 330 pp, $10. Contains an excellent selection of topics designed to fill the needs of the terminal student in mathematics. Topics include intuitive calculus and finite probability. K.W.

GENERAL, PSYCHOLOGY, E, S, B, L, *An Introduction to Piaget.* P.G Richmond. Basic Books, 1971, 120 pp, $4.95. Neither a development nor an exegisis this book provides a very readable introduction to and bibliography for the concepts and writings of Piaget. Doesn't have much application at the college level. J.A.S.

HARMONIC ANALYSIS, P, L, *Lecture Notes in Mathematics-162: Harmonic Analysis on Reductive p-adic Groups.* Harish-Chandra. Notes by: G. vanDijk. Springer-Verlag, 1970, iv + 125 pp, $3.50 (P). Lectures given at The Institute of Advanced Study, Fall 1969, meant to "strengthen the Lefschetz principle, which, in the context of reductive groups, says that whatever is true for real groups is also true for p-adic groups." "These lectures may also be regarded as an attempt to justify a claim about the philosophy of cusp forms made some years ago." (See "Eisenstein series over finite fields," Stone Jubilee volume, Springer.) R.B.K.

HARMONIC ANALYSIS, P, L, *Lecture Notes in Mathematics-166: PGL_2 over the p-adics: its Representations, Spherical Functions, and Fourier Analysis.* Allan J. Silberger. Springer-Verlag, 1970, vi + 204 pp, $5.30 (P). A continuation of Mautner's work on spherical functions over the p-adics. The main contribution lies in the explicit description of the representations of PGL(2,$\underline{0}$) in terms of the restrictions of irreducible representations of PGL(2,Ω), where Ω is a commutative p-field and $\underline{0}$ is the ring of integers in Ω. R.B.K.

NUMERICAL ANALYSIS AND APPLICATIONS, T*(17: 2), P, L, *Mathematical Optimization and Economic Theory.* Michael D. Intriligator. P-H, 1971, xix + 508 pp, $13.95. Distinctive in covering both programming and control theory, this book is an introduction to and survey of static and dynamic optimization techniques and their applications to economic theory. Static optimization: classical, nonlinear and linear programming, game theory. Applications: the household, the firm, general equilibrium, welfare economics. Dynamic optimization: calculus of variations, dynamic programming, maximum principle, differential games. Application: optimal economic growth. Appendices on analysis and matrices. Contains challenging problems and useful footnotes and bibliographies. R.B.K.

PROBABILITY AND STATISTICS, P, L, *Lecture Notes in Mathematics-190: Martingales. A Report on a Meeting at Oberwolfach May 17-23, 1970.* Ed: Hermann Dinges. Springer-Verlag, 1971, 75 pp, $3.50 (P). The intention of the Oberwolfach conference was that the following

question be discussed: "Martingales having proved to be a powerful tool, can they now be the subject of a probabilistic theory?" Speakers at the conference included D.L. Burkholder, H. Dinges, J.L. Doob, W. Hansen, F.B. Knight, P.A. Meyer, H. Rost, L.J. Snell, D.W.Stroock and S.R.S. Varadhan, and their talks are excerpted here. D.F.A.

PROBABILITY AND STATISTICS, T(18: 2), P, *Markov Chains.* David Freedman. Holden-Day, 1971, xiv + 382 pp, $18.95. Part of a trilogy on Markov processes, this text gives a sophisticated treatment of Markov chains in both the discrete and continuous time cases. The author's personal style of writing and sense of humor will delight many (and annoy others). No exercises, good bibliography. R.S.K.

PROBABILITY AND STATISTICS, T?, S, P?, *Statistical and Computational Methods in Data Analysis.* Siegmund Brandt. North-Holland, 1970, xii + 322 pp, $16.50. Intended for both students and research workers the book falls somewhere in between and satisfies neither. As a text it has no exercises and a rather uninteresting style of presentation. As a research aid it is incomplete. However, it does contain some redeeming features, such as its use of matrix algebra and examples of FORTRAN programs and printouts. R.S.K.

PROBABILITY AND STATISTICS, P*, L*, *Random Counts in Scientific Work.* Ed: G.P. Patil. 3 volumes, Penn St U Pr, 1970: *Random Counts in Models and Structures*, 268 pp, $11.50 (TR, December, 1970); *Random Counts in Biomedical and Social Sciences*, 267 pp, $11.50; *Random Counts in Physical Science, Geoscience, and Business*, 232 pp, $11.50. Contains all twenty papers given at the Biometric Society Symposium held in Dallas, Texas, in December, 1968, plus twenty-two additional invited papers. Volumes 2 and 3 give many interesting examples of the use of discrete distributions in real-life situations. R.S.K.

PROBABILITY AND STATISTICS, T*(15: 2), L, *Introduction to Mathematical Statistics, 4th Edition.* Paul G. Hoel. Wiley, 1971, x + 409 pp, $11.50. Although the material has been reorganized so that the first part can be used for a term course in probability, it is still primarily a good classical statistics text. R.S.K.

PROBABILITY AND STATISTICS, T(18), S, P, L*, *An Introduction to Probability Theory and Its Applications, Volume II.* William Feller. Wiley, 1971, 2nd ed, xxiv + 669 pp, $15.95. The topics covered are virtually the same as in the first edition, but the book has been rewritten "to make the reading easier." Although published posthumously, the revision had been completed at the time of Feller's death. R.S.K.

PROBABILITY AND STATISTICS, T(16-17: 1, 2), S, P, L, *Die Grundlehren der Mathematischen Wissenschaften, Band 172. Mathematical Methods in Risk Theory.* Hans Bühlmann. Springer-Verlag, 1970, xii + 210 pp, $15. "Attempts to create a synthesis out of a selection of modern scientific publications in the field of actuarial mathematics, with the goal of presenting a unified system of thought." Includes a brief review of probability theory with measure theory. Emphasizes model building--not parameter estimation. Covers applications to premium calculation, retentions and reserves, and stability criteria. F.L.W.

PROBABILITY AND STATISTICS, T**(15-16: 1), S, L*. *Introduction to Probability Theory.* Paul G. Hoel, Sidney C. Port and Charles J. Stone. Houghton Mifflin, 1971, xi + 258 pp, $10.95. First volume in a series of three on probability, statistics and stochastic processes. A well-written introductory text with particularly good sections on expectation. Also contains some good optional material, such as information on the characteristic function and a short chapter on random walks and Poisson processes. Good problem sets--answers included. R.S.K.

PROBABILITY AND STATISTICS, T*(16: 1), L*. *Introduction to Statistical Theory.* Paul G. Hoel, Sidney C. Port and Charles J. Stone. Houghton Mifflin, 1971, xi + 237 pp, $10.95. Second volume in a series of three on probability, statistics and stochastic processes. A well-written text, but requiring more sophistication than the first volume. Based on a decision theoretic approach, it includes optional sections on Bayesian methods. Contains material on testing the general linear hypothesis and an optional chapter on nonparametric methods. Omits concept of sufficiency. Good problem sets--answers included. Tables. R.S.K.

REAL ANALYSIS, FUNCTIONAL ANALYSIS, P, L. *Abelian Subalgebras of Von Neumann Algebras. Memoirs of the American Mathematical Society, Number 110.* Donald Bures. AMS, 1971, 127 pp, $2.20 (P).

JOURNALS

General Topology and its Applications, Volume I, 1971. Ed: S.P. Franklin. North-Holland, $20 per year ($13 for an individual associated with a subscribing institution). Published quarterly the journal defines general topology flexibly and currently includes "the axiomatic, set theoretic and geometric facets of topology as well as areas of interactions between general topology and other mathematical disciplines, e.g. topological algebra, topological dynamics, functional analysis, category theory, etc." J.A.S.

The Two-year College Mathematics Journal, Volume I, 1970. Ed: Joseph Hashisaki. Prindle, $20 per year (two copies of each issue), $5 for individuals. Published Fall and Spring. Nice format, more "education" than mathematics; the book reviews (more correctly book descriptions) are telegraphic and written by the publishers! The concerns are remedial, or as the journal prefers, "developmental." There is real need for a journal at this level to deal with problems in education and mathematics that are psychologically and mathematically significant. This journal does not do that; rather, it reinforces the view that both mathematics and the teaching of mathematics are discrete sets of isolated techniques. J.A.S.

Reviewers Whose Initials Appear Above

David F. Appleyard, Carleton; Lorraine L. Keller, St. Olaf; Roger B. Kirchner, Carleton; Richard S. Kleber, St. Olaf; Loren C. Larson, St. Olaf; J. Arthur Seebach, Jr., St. Olaf; Kenneth Wegner, Carleton; Frank L. Wolf, Carleton.

NEWS AND NOTICES

EDITED BY RAOUL HAILPERN, SUNY at Buffalo

Readers are invited to contribute to the general interest of this Department by sending news items to Mathematical Association of America, 1225 Connecticut Avenue, NW, Washington, D. C. 20036. Items must be submitted at least two months before publication can take place.

PERSONAL ITEMS

Professor Emeritus H. E. Bowie, American International College, received an honorary Doctor of Science degree from AIC on May 30, 1971.

Professor D. S. Ray, Bucknell University, represented the Association at the inauguration of H. J. Burnett as President of Washington and Jefferson College on April 3, 1971.

Bucknell University: Mr. David Finkel, University of Chicago, has been appointed Assistant Professor; Assistant Professor E. M. Luks has been promoted to Associate Professor; Assistant Professor P. I. Nelson, Rutgers University, has been appointed Assistant Professor; Mr. P. L. Strong, University of Illinois, has been appointed Assistant Professor.

University of Hawaii: Professor Adolf Mader is on sabbatical leave during the academic year 1971–72, when he will be at the University of Tubingen; Dr. Dale Myers, University of California, Berkeley, has been appointed Assistant Professor; Dr. Ruth E. M. Wong has been appointed Associate Chairman of the Department of Mathematics.

Indiana University, Bloomington: Associate Professor Peter Fillmore has been promoted to Professor; Assistant Professor Daniel Maki has been promoted to Associate Professor.

Indiana University-Purdue University at Indianapolis: Drs. Elaine V. Alton and J. E. Kuczkowski have been promoted to Associate Professors.

Oakland University: Professor John Dettman has returned from a leave of absence during which he was in residence at the University of Glasgow, Scotland; Dr. G. F. Feeman has been appointed Acting Chairman of the Department of Mathematics; Assistant Professor Jon Froemke has been promoted to Associate Professor; Mr. Kent Westerbeck, Case Institute of Technology, has been appointed Visiting Assistant Professor.

San Jose State College: Professor M. T. Bird retired on July 1, 1971 with the title of Professor Emeritus; Assistant Professor F. B. Fuller has been promoted to Associate Professor.

Texas A and M University: Drs. A. M. Hobbs, University of Waterloo, and W. L. Perry, University of Illinois, Urbana, have been appointed Assistant Professors; Associate Professor A. H. Stroud, SUNY at Buffalo, has been appointed Professor.

Worcester Polytechnic Institute: Professor J. J. Malone, Jr., Texas A and M University, has been appointed Professor and Head of the Mathematics Department; Professor and Acting Head of the Mathematics Department J. P. van Alstyne has been appointed Dean of Academic Advising; Mr. David Fraser, Brown University, has been appointed Assistant Professor; Assistant Professor Bernard Howard has been promoted to Associate Professor; Assistant Professor B. C. McQuarrie has been promoted to Associate Professor and is the first recipient of the Harold J. Gay Chair in Mathematical Science.

Dr. E. E. Blanche, Data Processing Consultant and former Chief Statistician from the Army's Logistics Division, has been appointed Acting President of the Capitol Institute of Technology, Kensington.

Assistant Professor L. E. Bragg, University of Kentucky, has been appointed Assistant Professor at the Florida Institute of Technology, Melbourne.

Professor Michael Capobianco, St. John's University, has been appointed Chairman of the Division of Natural Science at Notre Dame College of St. John's University, Staten Island.

Mr. Carlos Fallon, Manager, Value Analysis, RCA Corporate Staff, has been elected and installed National President of the American Society of Value Engineers.

Dr. S. T. Kao, University of New Mexico, has been named Acting Chairman of the Department of Mathematics and Statistics.

Associate Professor Lola F. Kiser, Birmingham-Southern College, has been promoted to Professor.

Associate Professor Edward Miranda, St. John's University, Jamaica, New York, has been appointed Chairman of the Mathematics Department.

Assistant Professor A. C. Segal, University of Alabama, Birmingham, has been promoted to Associate Professor.

Dr. W. L. Waltmann, Wartburg College, has been appointed Chairman of the Department of Mathematics.

Professor W. O. Alexander, University of Corpus Christi, died on April 14, 1971 at the age of 44. He was a member of the Association for sixteen years.

Professor Emeritus C. A. Hutchinson, University of Colorado, died on January 13, 1970 at the age of 72. He was a Charter Member of the Association.

Professor Emeritus A. B. Mewborn, Naval Postgraduate School, died on April 24, 1971 at the age of 67. He was a member of the Association for thirty-nine years.

Dr. J. T. Rosenbaum, University of Pittsburgh, died on June 22, 1971 at the age of 35. He was a member of the Association for seven years.

FELLOWSHIP AND RESEARCH OPPORTUNITIES IN THE MATHEMATICAL SCIENCES

In its annual brochure on Fellowship and Research Opportunities in the Mathematical Sciences, the Division of Mathematical Sciences of the National Research Council calls attention to a number of fellowships and other kinds of support for research in the mathematical sciences at both the predoctoral and postdoctoral levels to be awarded during the year 1971–72. Copies of this brochure are available from: Division of Mathematical Sciences, National Research Council, 2101 Constitution Avenue, N.W., Washington, D. C. 20418.

MATHEMATICAL ASSOCIATION OF AMERICA

Official Reports and Communications

THE FIFTY-SECOND SUMMER MEETING OF THE ASSOCIATION

The Fifty-second Summer Meeting of the Mathematical Association of America was held at Pennsylvania State University, University Park, Pennsylvania, from Monday, August 30, to Wednesday, September 1, 1971, in conjunction with meetings of the American Mathematical Society, the Pi Mu Epsilon Fraternity, and Mu Alpha Theta. There were registered 910 persons, including 683 members of the Association.

Sessions of the Association were held on Monday morning and afternoon, on Tuesday morning, and on Wednesday afternoon. All sessions were held in Schwab Auditorium. Presiding officers at the three Earle Raymond Hedrick Lectures were President Victor Klee, First Vice-President Garrett Birkhoff, and Second Vice-President J. W. Jewett; at the panel discussion on Monday morning, Professor Paul Axt; at the panel discussion on Monday afternoon, Professor Christine W. Ayoub; at the lecture by Professor R. L. Wilder on Tuesday morning, Professor R. G. D. Ayoub; and at the first panel discussion

on Wednesday afternoon, Professor J. S. Mamelak, and at the second, Professor C. B. Allendoerfer. The twentieth series of Earle Raymond Hedrick Lectures was delivered by Professor Abraham Robinson of Yale University.

The Program Committee consisted of R. G. D. Ayoub, Chairman; Steve Armentrout, P. C. Hammer, R. E. Huff, H. L. Krall, and J. S. Mamelak.

FIRST SESSION OF THE ASSOCIATION

Welcome on behalf of the University by Professor R. G. D. Ayoub of Pennsylvania State University.

The Earle Raymond Hedrick Lectures: *Nonstandard Analysis and Nonstandard Arithmetic*, Lecture I, by Professor Abraham Robinson, Yale University.

Nonstandard Analysis is based on a system which, in addition to the ordinary reals, contains also numbers that are infinitely small or infinitely large. In most other respects the system is quite similar to the real numbers, and this permits a consistent development of the calculus in which, for example, the derivative can be defined in terms of infinitely small increments of the variables. Corresponding ideas can be developed for other mathematical theories, e.g., in point set topology, in functional analysis, and in algebraic number theory. The lectures included nonstandard proofs of familiar results as well as of theorems which were first established by these methods.

Panel Discussion: What Undergraduate Courses Will be Taught in 1984?—A Look into the Future

A panel discussion with Professor Garrett Birkhoff, Harvard University, Professor Murray Gerstenhaber, University of Pennsylvania, and Professor J. B. Rosser, University of Wisconsin, Madison, and Mathematics Research Center.

Professor Birkhoff spoke on "Computers and Future Undergraduate Mathematics Education." A great variety of mathematical courses should be available to undergraduates in 1984, and many of these should involve the use of computers, both to reduce drudgery and to increase the precision and sophistication of thinking.

Thus the meanings of approximation and convergence can be made much more vivid, and the fundamental theorem of the calculus more convincing, by enabling the student to carry out algorithms requiring a large number of steps (10^3–10^6, say). Moreover, this capability makes the calculus an even more powerful and impressive tool.

The computer also frees courses in real and complex algebra from its traditional humiliating practical limitation to quadratic (or perhaps cubic) polynomial equations, and linear systems in two (or possibly three) unknowns. In the area of symbolic algebra, it suggests a host of fundamental new problems, and vistas into logic and "mechanical thinking" which we have only begun to explore. Statistical data-processing constitutes another fascinating area whose scope and scale have been enormously extended by computers. Experiments using "random numbers" should also make courses on probability much more interesting.

But the greatest opportunities for undergraduate intellectual development, as of 1984, may concern tutorial work on individual projects using powerful but inexpensive computers. For scientifically minded students, indeed, a fertile source of undergraduate thesis topics is already provided by *scientific computing*.

Professor Birkhoff discussed the above possibilities in some detail, basing his comments on personal experience where possible.

Professor Gerstenhaber suggested that programming, computational techniques in linear algebra, probability, and statistics will, by 1984, compete strongly with the traditional calculus for the limited time available in the first two years of undergraduate mathematics training. These topics will be taught in integrated courses in which proofs are kept at a minimum, rigor being reserved for "honors" sections. These and other "practical" topics will be amplified in junior and senior level courses for mathematics majors, a majority of whom will be women going into industry. Research-oriented students will be transferred to graduate courses as quickly as possible.

Professor Rosser felt that, if the past trends continue to 1984, the teaching in mathematics courses of actual algorithms will be nearly non-existent. Meanwhile, more and more freshmen will

enter college with computer backgrounds. The pressures to have students taught algorithmic skills are increasing. The result could well be that by 1984 most of the present clientele for freshman and sophomore mathematics courses would be diverted to computer science courses. This is not desirable because basic understanding of the algorithms is a mathematical attainment. To avert such a diversion, the basic freshman and sophomore courses (and a fair number of higher courses) must feature algorithms attractively and competently, especially with attention to their possible applications.

This was followed by a general discussion by the panel and the audience.

SECOND SESSION OF THE ASSOCIATION

Hedrick Lecture II, by Professor Robinson

Panel Discussion: Women in Mathematics

A panel discussion with Professor Mary Gray of American University, Professor Gloria C. Hewitt, University of Montana, and Professor Mary E. Rudin, University of Wisconsin, Madison, with Professor Christine W. Ayoub as moderator.

Professor Ayoub suggested that the panel center their discussion around the following questions:

1. Are there a substantial number of women with mathematical potential who get discouraged somewhere along the line? If so, when, why and how? Is mathematics different from other disciplines in this respect?
2. Assuming that the answer to the first part of (1) is yes, should an effort be made to see that these women are encouraged—or at least, not discouraged more than is appropriate in the present situation? If so, how should this be done?
3. How much discrimination do women encounter at various stages in their careers?
4. What measures can be taken to overcome this discrimination—and to make it possible for women to have a successful mathematical career (even if they get married and have children)?

Professor Rudin felt that the development of mathematical talent is highly cultural, and that women are certainly a mathematically culturally deprived group. However, there is little overt discrimination except on the new Ph.D. who may suffer in job opportunities because of the high drop out rate of her sisters. A real boon in the use of talent and training would be more part-time jobs.

Professor Hewitt stated that no one can deny the existence of discriminatory practices against women in mathematics. Too often in making decisions to hire women, marital status, family responsibilities, family size, and such are influential factors. Neopotism rules are invoked, or invented, to justify marginal appointments without fringe benefits or for rejecting the applicant, whereas it should only be assumed that the applicant will desire regular, full-time, permanent employment unless there is clear evidence to the contrary.

Recommendations accompanying applications for women often support the myth that women are a poor risk. There are those which declare that the applicant is one of the best students he has ever had, truly exceptional—for a woman she should excel in mathematics.

For men and women of equivalent standing, there are too often large discrepancies in salaries, fringe benefits, departmental duties, appointments to key departmental committees. Promotions are much slower for women than for men. The criteria should be the same.

Professor Gray suggested that, rather than decrying the discrimination against women that exists, the Association of Women Mathematicians proposes to do something to overcome it and to encourage women mathematicians. The AMS and MAA can help improve the image of women-as-mathematicians by including women in their leadership, by featuring women in films, etc. Female students and young researchers should be conditioned to think of themselves as potential Fields medal winners and then should get some encouragement—e.g., a few Sloan fellowships. The overt discrimination is being combatted by HEW compliance and other actions; it is the more subtle forms that need some attention.

This was followed by a general discussion by the panel and the audience.

THIRD SESSION OF THE ASSOCIATION

Hedrick Lecture III, by Professor Robinson

Business Meeting of the Association; presentation of Lester R. Ford Awards.

History in the Mathematics Curriculum: Its Status, Quality, and Function, by Professor R. L. Wilder, University of Michigan and University of California, Santa Barbara.

Student complaints regarding the great abstractness and seeming unrelatedness of their courses suggests the need for a course which would broaden their perspective and improve their understanding of modern mathematics. A course which explains how and why mathematics evolves seems indicated. This would not be a history course of the traditional type, but would use the facts of history as a basis for investigating the manner in which mathematical concepts and structures come to be created, and clarify the way greater abstraction has permitted consolidation and simplification of seemingly unrelated and complex theories.

FOURTH SESSION OF THE ASSOCIATION

Panel Discussion: Concerns of Community Colleges

A panel discussion with Professors C. A. Lathan, Monroe Community College, Rochester, New York, and Professor Ralph Mansfield, Loop College, Chicago, with Professor J. S. Mamelak, Community College of Philadelphia, as moderator.

Professor Mamelak opened the discussion by stressing his belief that the central experience in the community college is the need to fractionalize all existing instruments of instructions. A decomposition of courses, examination procedures, entrance requirements, etc. into as many parts as are required to accommodate the increasingly large and heterogeneous body of students is clearly necessary. The challenge in the community college is to find the best means of fractionalizing the existing instruments of instruction and evaluation to achieve the goals set for the community college.

Community colleges are asked to accommodate an increasingly large body of students of highly varied interests, background and motivation, in a social climate which does not permit labelling of individuals as academic failures and a psychological setting in which authority of the subject, professor or policy generates little, if not negative response. The process of adopting existing academic instruments to the existing environment within the community college and the steps necessary to move the academic community to accommodate its practices to this reality were discussed under the headings of: 1) Remediation (Professor Lathan); 2) Course content (Professor Mansfield); 3) Personnel and transcript transfer problems (Professor Mamelak).

Professor Mansfield observed that community college curricula in mathematics are fragmented by the various technology groups demanding special courses in mathematics and the diverse abilities of students who take such courses. These non-mathematician courses fail to develop student abilities to translate real situations into mathematical terms.

It was suggested that mathematics course content for non-mathematicians should develop student abilities to use mathematics as a language and deal successfully with mathematical models that can be solved and translated into real situations. This can best be accomplished by combining the teaching skills of mathematicians with those of technical specialists in team teaching efforts.

Professor Lathan stressed that it was important to realize that there is a sizable student population in both two-year and four-year colleges taking courses with titles such as college arithmetic, elementary algebra, intermediate algebra, and introduction to college mathematics. Recently CUPM (Committee on the Undergraduate Program in Mathematics) published a proposal regarding *A Course in Basic Mathematics for Colleges*. The speaker responded to this proposal by offering both positive and negative aspects of the report.

Professor Mamelak felt that Ph.D.'s are needed at community colleges provided they are willing to sensitize themselves to the pedagogical problems at these institutions. The problem of grading and transcript evaluation has become very complex since the requirements of the academic community and the purposes of the community college are frequently contradictory.

This was followed by a general discussion by the panel and the audience.

Panel Discussion: Placement Tests in Mathematics—How Valid Are They?

A panel discussion with Mr. Gene Murrow, Harvard School, North Hollywood, California, Professor Alex Rosenberg, Cornell University, and Dr. Marion G. Epstein, Educational Testing Service, Princeton.

Mr. Murrow expressed certain concerns regarding the value of placement tests in mathematics and suggested that oral examinations might be more appropriate for proper placement of students in college calculus courses.

Professor Rosenberg gave some facts and figures on the college experiences with Advanced Placement in Calculus. He provided a detailed description of the experiences with Advanced Placement Freshmen in Mathematics at Cornell University.

Dr. Epstein defined the different types of test validity with which a test developer or test user might be concerned. It included a brief discussion of the methodology of predictive validity studies and of the interpretation of validity coefficients. The limitations on the use for placement of tests designed primarily for selection were considered, and the types of tests valid for vertical or advanced placement were contrasted with those suitable for use in horizontal sectioning, with examples from national testing programs.

This was followed by a general discussion by the panel and the audience.

SPECIAL SESSIONS OF THE ASSOCIATION

Film showings were held in the Auditorium of the Conference Center on Sunday, Monday, and Tuesday at 7:00 P.M. The showings included several films not previously shown in public. The following films were shown on Sunday:

Films of the MAA Individual Lectures Film Project (ILFP)

7:00–7:25 P.M. SHAPES OF THE FUTURE I—SOME UNSOLVED PROBLEMS IN GEOMETRY—TWO DIMENSIONS, with Victor Klee (in color)

7:35–8:15 P.M. SHAPES OF THE FUTURE II—SOME UNSOLVED PROBLEMS IN GEOMETRY—THREE DIMENSIONS, with Victor Klee (in color)

A Film of the Educational Broadcasting Corporation of New York

8:25–8:54 P.M. NEW WORLD, NEW MATH (in color)

Among the films shown on Monday were the following:

Films of the NCTM Series: Elementary Mathematics for Teachers and Students (in color)

7:00–7:11 P.M. BETWEEN RATIONAL NUMBERS (KNIGHTS)
7:12–7:22 P.M. RECIPROCALS—MULTIPLICATIVE INVERSES (SUNGLASSES)
7:23–7:33 P.M. THE BIGGEST RECTANGLE
7:34–7:44 P.M. HIDDEN TREASURE
7:45–7:56 P.M. EXPLOITATION OF ERRORS (EDGAR'S GUESS)
7:57–8:07 P.M. SOLVING PAIRS OF EQUATIONS (PIRATES)
8:18–8:30 P.M. PROBABILITY (RAJAH)

The film GRAPHING INEQUALITIES (MARVELOUS MARSHES) scheduled for showing from 8:08 to 8:17 P.M., was not shown since it was not received from the distributor.

Among the films shown on Tuesday were the following:

Allendoerfer Geometry Films (animated and in color)

7:00–7:10 P.M. GEOMETRIC CONCEPTS
7:11–7:21 P.M. AREA AND PI
7:22–7:32 P.M. GEOMETRIC TRANSFORMATIONS

In addition, on Monday and Tuesday evenings films of the MAA MATHEMATICS TODAY series were shown.

MEETING OF THE BOARD OF GOVERNORS

The Board of Governors of the Association met on Sunday at 9:00 A.M. in Rooms 402–403 of the Conference Center with thirty-four members present. Among the items of business transacted were the following:

In its belief that the effectiveness of the MAA in fulfilling its purposes is heavily dependent upon the efforts of its individual members working through the local Sections and that, especially at the present time, it is important for all Sections to be organized for maximum effectiveness, the Board voted that all Sections be requested to review their By-Laws by a committee of the Section, in accordance with the suggestions made at the meeting of Section Officers on August 24, 1970, at the University of Wyoming, and contained in the brochure "Guidelines for Sections" (available from the Washington office of the MAA). Particular attention is directed to suggestion 6, namely that, since it takes a year to learn the routine of the office of Section Chairman, it is unwise to have him serve merely a year and that, accordingly, it is recommended that Section By-Laws provide for a Chairman to serve first for one year as Chairman-Elect, then for two years as Chairman, and that he remain on the Executive Committee for an additional year as Immediate Past-Chairman.

Attention is also called to suggestion 4, namely that, because of his key role, a Secretary should be elected for a term of at least three years and should have the option of being renominated and reelected.

As suggested in items 7 and 8, each Section is urged to have a reasonable number of officers committed to its welfare, for example, by the appointment of a First and Second Vice-Chairman, so that the Executive Committee will consist of at least six persons.

It was emphasized that Sections are in no way required to conform precisely to these suggestions, as local situations may suggest other arrangements of equal effectiveness. Thus, some Sections have arrangements whereby a Vice-Chairman automatically becomes Chairman after expiration of his term, in which case the suggested one-year service of a prospective Chairman as Chairman-Elect might not be needed. It was suggested, however, that, after Sections have reviewed their By-Laws in accordance with the above-mentioned brochure and when they submit the revised By-Laws to the Committee on Sections for approval, a statement of explanation be included wherever such revised By-Laws differ significantly from those suggested in the brochure, in order that the Committee is informed of the thinking of the Section which causes any possible discrepancies.

The Board of Governors considered at great length the serious prospect that NSF support of a separate CUPM Central Office and for meetings of the Commission may be unavailable beyond the middle of 1973. After considerable debate and numerous changes, the following resolution was approved by the Board:

"RESOLVED THAT:

1. Since new problems have arisen in mathematical education, and new methods of attack may be required to deal with them, the Association requests its President to appoint an ad hoc Committee on New Priorities for Undergraduate Education in the Mathematical Sciences. This Committee should be asked to identify the most important problems and to recommend initial steps leading toward their solution.

2. The Association expresses pride in the achievements of its Committee on the Undergraduate Program in Mathematics (CUPM) and a strong belief in the current importance of problems now being attacked by CUPM (technical-occupational mathematics, the role of computing in the study of mathematics, alternatives to current freshmen mathematics programs for the general student, special problems of minority groups, mathematics for the social sciences, applied mathematics, teacher training). In order to bring these attacks to fruition and to avoid wasting funds already committed,

continued Federal support may well be required until at least the middle of 1974 for the individual CUPM projects, for operation of the CUPM Central Office, and for meetings of the Commission.

3. Since direct Federal support, for the indefinite future, of a separate CUPM Central Office and of meetings of the Commission appears to be unlikely, an urgent initial charge for the Committee on New Priorities should be the making of provisions for an orderly transition from CUPM's activities to those of a new mechanism for attack on educational problems. This will insure a continuity of effort in those areas in which the new problems identified by the Committee are related to the ones currently being attacked by CUPM. The Committee should give special consideration to the desirability of having all future MAA projects managed by an MAA Projects Director from the MAA Central Office in Washington, D. C.

4. The Committee on New Priorities should devote special attention to the sort of problems raised in the recent Carnegie Commission report "Less Time, More Options," and to the following specific questions:

a) How is mathematics involved, and how should it be involved, in current and likely future national concerns, such as population, pollution and environmental control, transportation, etc?

b) How can mathematics be taught more effectively in "service courses" for those wanting to apply it in other disciplines, and how can mathematicians be better prepared to teach such courses?

c) What are the important job opportunities, outside of mathematics, for those with mathematical training, and how should that training be improved so as to better fit them for such jobs?

5. The Association recognizes that all of the constituent organizations of the Conference Board of the Mathematical Sciences have a vital interst in undergraduate mathematics education, and that several of them have their own committees on education. In view of this broad base of professional concern, the Committee on New Priorities should consider possibilities for cooperation with the CBMS Committee on Education in the Mathematical Sciences. It should also consider the advisability of the following (and if they are found to be advisable, should propose specific times and agenda):

a) A public meeting, open to all mathematicians, to discuss new directions and new priorities for undergraduate education in the mathematical sciences. (This might be held at a winter or summer meeting of MAA-AMS-CBMS.)

b) A conference dealing with the same topic, to which representatives of all national organizations in the mathematical sciences would be invited.

c) Alternative meetings, conferences, or symposia."

The announced intention of NSF to discontinue support for the CUPM Central Office and meetings of the Commission resulted in a discussion of the current role the mathematical community is playing in decisions of NSF affecting the mathematical community. There was considerable dissatisfaction with the fact that priorities for spending in Federal support of mathematics are too frequently set without adequate consultation with the mathematical community. The following resolution was approved:

"In its relatively short history, the National Science Foundation (NSF) has been a vital force in the development of science, in general, and of the mathematical sciences, in particular, for the general national welfare. However, in recent years, there have been policy decisions by the NSF, affecting the mathematical sciences community (and other scientific disciplines), that seem to have been taken without adequate consultation with the community involved. In our own area, we can cite the plight of the Committee on the Undergraduate Program in Mathematics (CUPM), that of the Advanced Science Seminar Program, and the NSF intervention in the nature of the National Information

System in the Mathematical Sciences, and the sharp cut-backs in fellowship support.

The task of the NSF is to "promote research and education in the sciences." In our view, the proper attainment of this goal requires close cooperation and consultation, on the part of the NSF, with the scientific community, in the formulation of the appropriate policies and mechanisms.

It is resolved that the Board of Governors of the MAA urge upon NSF a re-evaluation of the relationship of NSF to the mathematical community, with an eye towards a larger involvement by that community in NSF decisions affecting it.

It is further resolved that this resolution be transmitted to the appropriate officials of the Federal Government and to the other organizations in the mathematical sciences."

The Board considered the final report of the ad hoc Committee to Consider Certification and Accreditation in Mathematics. It approved all of the Committee's recommendations including the recommendation that the Association take no action toward establishing a system of accreditation or certification in mathematics at this time. The report and its recommendations will appear in an early issue of this MONTHLY.

The Board also voted to request the President to appoint a committee to attempt to set up guidelines for the evaluation of mathematics departments and undergraduate programs: such guidelines could be distributed to the six regional commissions for general accreditation of institutions of higher learning in the United States. Analogous guidelines for the mathematics programs of precollege teachers could be distributed to the National Council for Accrediting of Teacher Education (NCATE) for their use in accrediting college programs of teacher training. In addition, such guidelines could be used by individual departments for voluntary self-evaluation. The committee was requested to submit these guidelines to the Board of Governors for consideration, possible approval and transmittal to the agencies mentioned above.

The Board voted to amend its action taken a year ago concerning the establishment of three types of corporate memberships, by setting the dues for "Corporate Members" at $200 (rather than $300), with the dues for "Sustaining Corporate Members" and "Sponsoring Corporate Members" to remain at $500 and $1000, respectively.

The Executive Director reported the membership of the Association as 17,961 individual members, almost unchanged from a year ago, 3 corporate members, and 340 academic members, the latter an increase of 8.2% over the corresponding figure a year earlier.

The Board voted that the ad hoc Committee on the Role of the Two-Year College Teachers of Mathematics in the Association and the standing Committee on Assistance to Sections on Two-Year College Problems be replaced by a new standing Committee on Two-Year College Teachers of Mathematics.

BUSINESS MEETING OF THE ASSOCIATION

A business meeting was held on Tuesday morning with President Klee presiding.

The seventh set of Lester R. Ford Awards was presented by President Klee to authors of expository articles published in the MONTHLY and MATHEMATICS MAGAZINE in 1970. The Awards, in the amount of $100 each, were presented for six articles (for further details on these Awards, see the August–September issue of this MONTHLY, page 830).

President Klee discussed the serious prospect that NSF support for a separate CUPM Central Office and meetings of the Commission may be unavailable after June 30, 1973. This would presumably necessitate a change in the Association's procedures for attacking educational problems in the mathematical sciences. He then read the two resolutions concerning this matter passed by the Board of Governors at its meeting on August 29, 1971.

The Secretary reported on some of the actions taken by the Board of Governors since its meeting in January. He announced the election of Professor E. S. Langford as Associate Editor of the MONTHLY in charge of the University of Maine Problems

Group to succeed Professor G. P. Murphy, and Professor Jane W. Di Paola as an additional Associate Editor of the MONTHLY. Professor Di Paola will assume responsibility for technical editing of manuscripts and proof sheets. Professor Harry Pollard has been elected an Associate Editor of the MATHEMATICS MAGAZINE to succeed Professor E. A. Maier.

The Secretary noted that the major event during the current calendar year in the Washington office was the signing of a contract with Computer Sciences Corporation for the development and maintenance of a computerized membership and subscriber record system. The decision to sign a contract with Computer Sciences Corporation culminated almost eighteen months of study and negotiation, during which time eleven service bureaus and three computer manufacturers were considered. It is hoped that the result of the conversion to computerized operations will be a more rapid and more accurate processing of address changes.

The Secretary announced that the Association was very pleased to acknowledge the receipt of two grants: one for $28,250 from the National Science Foundation (NSF) for partial support of the Visiting Lecturers Program for 1971–72, and one for $5,000 from the grant by NSF to the Committee on a National Information System in the Mathematical Sciences (NISIMS) of the Conference Board of the Mathematical Sciences. Of the latter amount $3,000 was allocated for the further development by the MAA of plans for a classified cumulative index for the MONTHLY, to be carried out under the direction of Professor K. O. May, and $2,000 for a survey to determine the information desires of MAA members. This survey is currently being planned by an ad hoc Committee on a Survey of the Membership of the Association under the chairmanship of Professor E. F. Beckenbach. The questionnaire being prepared by this Committee will be mailed to the membership early in October.

The Secretary announced several new publications, including MAA STUDIES IN MATHEMATICS, Volume 7, "Studies in Applied Mathematics," edited by Professor A. H. Taub, which will appear in October, a Slaught Paper, "Differentiation of Integrals" by Professor A. M. Bruckner, to appear before the end of the year, and another Slaught Paper, tentatively titled "Nonstandard Analysis," to appear next year.

The following schedule of future meetings of the Association was announced: Las Vegas, Nevada, January 19–21, 1972; Dartmouth College, Hanover, New Hampshire, August 28–30, 1972; Dallas, Texas, January 27–29, 1973; University of Montana, Missoula, August 20–22, 1973; San Francisco, California, January 17–19, 1974; Washington, D. C., January 25–27, 1975; San Antonio, Texas, January 24–26, 1976.

The Secretary extended on behalf of the Association sincerest thanks to the local Committee on Arrangements for their splendid job in attending so meticulously to all details of the meeting. He expressed special appreciation to Professor D. C. Rung, Chairman of the Committee, for coordinating so effectively all aspects of the arrangements and for planning so energetically and thoughtfully for the various events scheduled for the meeting.

Professor Anatole Beck presented a resolution concerning the war in Southeast Asia. President Klee requested the Secretary to read that part of the By-Laws describing the purposes of the Association. President Klee then read the statement from Robert's Rules of Order that "a main motion that proposes action outside the scope of the organization's object as defined in the By-Laws or corporate charter is out of order unless the assembly by a two-thirds vote authorizes its introduction." He suggested to Professor Beck that he ask the meeting for authorization to consider his motion. Professor Beck felt that his motion was not out of order and appeared the ruling of the chair that it was. After a brief debate, the ruling of the chair was sustained by a large majority in a standing vote.

A further motion by Professor Beck that the meeting authorize consideration of his motion despite the fact that it was considered to be outside the Association's By-Laws failed to carry by a standing vote.

MEETING OF SECTION OFFICERS

The meeting of representatives of the Section was held on Monday evening in Rooms 402–403 of the Conference Center. Professor L. E. Mehlenbacher, Chairman of the Committee on Sections, presided. Fifty-nine persons were present representing twenty-six of the twenty-eight Sections of the Association.

President Klee welcomed the representatives of the Sections and guests. He stressed the importance of the work of the Sections and expressed the hope to be able to visit more Sections in the coming year.

The Executive Director, Dr. A. B. Willcox, reported on the new formula, authorized by the Board of Governors at its meeting on August 29, for reimbursing official delegates to the meeting of Section Officers. The new formula was designed to approximate the cost of round-trip tourist air fare between the delegate's home and the place of the meeting.

The Executive Director reported that only two proposals for grants from the Fund for Aid to Sections had been received in 1971. Only four awards had been made in 1970. As a result, the 1972 budget for the Fund has been reduced. The level of the Fund remains substantially above the maximum annual use since its establishment, and Sections were urged to make use of it, in accordance with the "Guidelines for Proposals" contained in the brochure "Guidelines for Sections".

In answer to a question as to what kinds of projects are supported from the Fund for Aid to Sections, the Executive Director gave as examples (1) speakers at meetings, (2) special projects to study the problem of interface between two-year and four-year colleges, and (3) a small visiting lecturer program to secondary schools.

The Executive Director announced that a display of MAA publications was available for use at Section meetings. Books and other publications may be seen, felt, and examined. They may be purchased by members attending Section meetings at the special members' prices. There will be an additional 10 percent discount for publications ordered at Section meetings. This additional saving may be passed on to the purchaser or kept for the Section treasury at the discretion of the Section. The exhibit was on display at the meeting of Section Officers. The Executive Director demonstrated the setting up of the display.

Professor H. M. Bacon, Chairman of the Committee on Secondary School Lecturers, led a discussion on "How to Organize Section-Sponsored Secondary School Lecturer Programs." He noted that the MAA Committee on Secondary School Lecturers hopes that, in the absence of financial support for a Secondary School Lecturer Program on a national scale, responsibility for funding and operating Section-sponsored lecturer programs will be assumed by many or all of the Sections. He then outlined briefly some of the aims of such programs. Organization of a program was discussed under the two principal headings of *Financing* and *Administration.*

Under Financing, a few possible sources of funds were listed under two main headings, namely, general funds belonging to the Section, and funds specially secured for the lecturer program. Under the latter there were noted (a) contributions from local business and industry, (b) contributions either of funds or of services by State or local Departments of Education, (c) contributions from high schools themselves, (d) funds for Section projects administered by the MAA national office, (e) cooperative arrangements with local or state associations of teachers of mathematics, (f) cooperative arrangements with State Academies of Science, (g) contribution of travel expenses by a lecturer's own institution.

Under Administration of a Section-sponsored program points were noted as follows. (1) A policy-setting committee of the Section would be useful. (2) A Director, probably chosen by the policy committee, is of great importance. He must have sound judgement, be interested and energetic, prompt and dependable in handling detail, tactful, reason-

ably well known, and a member of a college or university mathematics faculty. In a large Section, regional coordinators might assist the Director. (3) Suggestions were made concerning a "panel" of lecturers and about length of visits. (4) Lecture topics were discussed, and it was noted that experience seems to indicate that lectures in pure and applied mathematics or the history of mathematics are more successful than talks about such things as career opportunities, although this last kind of subject might well be taken up in informal conferences with individuals. (5) the question of honoraria and expenses of lecturers was touched upon, but it was noted that there will necessarily be a wide variation in practice because of the uncertainty of funding. (6) Publicity for a Section-sponsored program will vary from one Section to another, but the Association plans to make available a small brochure that can be used by all Sections as one aspect of an effort to make a program known. (7) The importance of continuity in direction of a program was stressed.

Professor L. L. Clarkson, Chairman of the CUPM Panel on Special Problems of Minority Groups, reported on the work of that Panel. The major thrust of this Panel has been directed at identifying and resolving mathematics and mathematics education problems in the Traditionally Black Institutions (TBI's) with the understanding that any useful information or techniques developed due to these efforts would have applications in a much wider range of institutions. In an attempt to make its efforts as meaningful as possible, the Panel maintains contact with other national panels, committees or organizations that are concerned, at least in part, with some of the same problems. To the speaker's knowledge, there are six such groups associated with major national professional mathematical organizations. He then surveyed briefly the relevant activities of these groups.

Professor K. W. Wegner, a member of the Association's Committee on Assistance to Developing Colleges (CADC) reviewed the work of that committee. Since its last meeting in January, the Committee has had the disappointment of the turning down of the proposal for the establishment of an Office of Awareness of Opportunities in the Mathematical Sciences. It had been the hope that this office would carry on in a big way some of the projects it was attempting in a very small way.

It is now hoped that many of the concerns of CADC will be taken over by the newly-formed National Association of Mathematicians made up of teachers at TBI's.

The change in the economic situation has made the previously active "employment register" less important. Other concerns of CADC overlap with those of the Panel on Special Problems of Minority Groups. Thus, some members of CADC wonder whether or not it should even continue to exist.

President Klee spoke on the concern about the future of CUPM. He discussed the serious prospect that NSF support of a separate CUPM Central Office may be unavailable after the middle of 1973. This would probably necessitate a change in the Association's procedures for attacking educational problems. The two resolutions passed by the Board of Governors at its meeting on the preceding day were read.

Professor Alex Rosenberg, Chairman of CUPM, felt that the situation regarding CUPM at the moment was extremely black. Support must come from the government or the MAA, and, if neither source can supply it, CUPM has to be scaled down drastically. Since no one knows how much the membership values CUPM, Professor Rosenberg suggested that the Sections might try to find this out; he would be interested in any information they could contribute to such an assessment. Small samples of opinion taken from the membership indicate that CUPM is highly regarded. In the eight years of its existence, over 300 people have been involved as members of the Commission or members of its panels; several thousand people have attended conferences sponsored by the Commission; the CUPM Central Office receives about 48,000 requests for various publications per year in addition to the numbers of documents it mails out automatically. He then summarized the activities of CUPM since January 1971 under the headings:

Panel on Teacher Training: The Panel met in February and in June to complete work on the revision of Recommendations and Course Guides for the training of elementary and secondary school teachers. These will be submitted to the Commission for approval at its August meeting.

Panel on College Teacher Preparation: The Panel met in February to continue work on a handbook for new instructors and to plan a conference on problems of handling large classes and on the training of teaching assistants. The conference was held on May 1. A Newsletter publicizing some of the ideas from the conference is planned for publication in the near future. The handbook should be completed at the next Panel meeting in October.

Panel on Mathematics in Two-Year Colleges: "A Course in Basic Mathematics," prepared by a group consisting of this Panel augmented with other two- and four-year college people, has been printed, and approximately 10,000 copies have been distributed. Two regional conferences (in New Orleans and Memphis) on this report have been held. The Panel is now turning its attention to the role of mathematics in technical-occupational curricula. An information-gathering conference was held in February. The Panel met to work on this problem again in April and July.

Panel on Computing: The report of the Panel entitled "Recommendations for an Undergraduate Program in Computational Mathematics" has been printed and is currently being distributed. A reconstituted Panel has turned its attention to the impact of the computer on mathematics courses. An information-gathering conference was held in January, and the group has met one time since then. Another meeting is planned for September.

Panel on Applied Mathematics: The Panel is working on options for the GCMC course Mathematics 10. Meetings were held in February and May. The next meeting will be in September, at which time the Panel should complete its report.

Panel on Statistics: The Panel's report "Preparation for Graduate Work in Statistics" has been printed and is being distributed. The Panel is now working on recommendations for a first course in statistics that would have no calculus prerequisite.

Panel on Special Problems of Minority Groups: The Panel met in March and June. Among its projects is the preparation of a handbook to be used by minority students in their selection of graduate schools and specific efforts to open the lines of communication between the Black and White mathematical communities.

Revision of GCMC: The group met in April. It is engaged in writing a Commentary on Mathematics 1–4 and Mathematics 6. The Chairman of the group is preparing a draft for the group to consider at its final meeting in October.

Panel on Innovations: This new Panel, chaired by Arnold Ross, with D. T. Finkbeiner, I. N. Herstein, Helmut Rohrl, and J. H. Wells as members, met in June. As an initial project, it is surveying departmental chairmen to determine what innovative techniques are being used which deserve wider publication, in a CUPM Newsletter for instance.

Consultants Bureau: By July 1, 1971, there had been 35 applications for consultant visits during the previous academic year. Twenty-nine of these visits were completed.

Additional Publication: "A Basic Library List for Two-Year Colleges" was printed in July and distributed to the Mathematics departments of two- and four-year colleges. The initial mailing was 4200 copies.

Professor Harley Flanders, Editor of this MONTHLY, announced that the MONTHLY is dependent upon suggestions from Section Officers for names of outstanding speakers at Section meetings, who might provide good articles for the MONTHLY. He suggested that names of such speakers be sent to him, along with names of other people in attendance at the meeting, so that a number of recommendations can be secured before the speaker is invited to submit an article.

Professor M. W. Pownall, Chairman of the Committee on Visiting Lecturers, re-

ported on the work of that Committee. During the past academic year, 169 colleges and universities which do not grant a Ph.D. degree in mathematics were visited under the auspices of the MAA Program of Visiting Lecturers, including 30 institutions not visited in the last five years. Although this represents a slight decline in total activity as compared with 1969–70 (when 183 colleges and universities were visited), the Committee is slowly building up a clientele among the two-year colleges and developing colleges. These colleges have not participated very much in the program in the past, and it is one of the Committee's major problems to increase participation from these institutions. The Committee will welcome the assistance of the Association and its Sections as it continues to work on this problem.

Professor J. M. Earl, Chairman of the Committee on High School Contests, reported on the 1971 Contest. He announced that this year the Charles T. Salkind Silver Cup was awarded to Freeport High School, New York, for the highest Team Score of 354.00 of a possible 450 points. A Small Plaque was awarded to James S. Pace of the same high school for the highest individual Score of 142.50 of 150 points.

Professor H. M. Cox, Executive Director of the Contest, announced that the date of the next Contest is March 14, 1972. He expressed appreciation to the Section Contest Chairmen and to the officers who have assisted with the Contest.

Professor S. L. Greitzer, Chairman of a Subcommittee on a USA Mathematical Olympiad, reported on the work of that subcommittee. Previous to February, 1971, the Committee on High School Contests had appointed a subcommittee to investigate the desirability of presenting a contest which would consist of very few problems whose solution would require mathematical maturity beyond that required for the Annual High School Mathematics Contest. This subcommittee apparently was dormant after the death of Professor Salkind. As the result of an article by Professor Nura Turner in the February issue of this MONTHLY, Professor Earl reactivated the subcommittee.

Correspondence with mathematicians and educators made it evident that many of them think there is a place for such a competition. At a meeting of the subcommittee on August 30, 1971, it was decided to recommend that such an "Olympiad" be instituted with a minimum of delay. It would be small, involving about 100 students who had earned top scores in the Annual High School Mathematics Contest, with about 8 of these selected as top scorers.

The help of the Section Officers might be needed, chiefly for publicity purposes. The recommendations will be submitted to the Committee on High School Contests on August 31. (Secretary's Note: The Committee approved the recommendation at its meeting on August 31, and the proposal for a USA Mathematical Olympiad has now been approved by the Board of Governors by a mail ballot. The first such Olympiad will be held in May 1972.)

Professor D. T. Finkbeiner, Chairman of the Committee to Consider Certification and Accreditation in Mathematics, reviewed the final report and recommendations of this Committee.

Professor L. C. Huffman, Midwestern University, reported on the Texas Section work with two-year colleges. This Section has for a good many years been trying to include the two-year colleges in a meaningful way. This effort can be summarized under four categories:

1. There is always at least one two-year college representative on the Executive Committee of the Section.

2. At least one session of the annual meeting is always designed especially for the two-year college teachers.

3. The annual meeting of the Section is held on the campus of a two-year college about every sixth year.

4. At the request of the two-year college representatives, a booklet is being written which will give the course description and textbook for the undergraduate mathematics

courses being taught in the various colleges and universities in the Texas Section.

Professor David Schneider, University of Maryland, reported on his experiences with handling large classes of students. He described in detail how a precalculus and probability course, which is intended for students majoring in the biological, social, and management sciences, is conducted at the University of Maryland. An enrollment of over 3,000 students per semester justified a major commitment by the mathematics department and resulted in a course which is beyond the capabilities of the standard classroom situation.

Students meet in groups of about 25, for 3 hours per week with a Graduate Assistant. The first 30 minutes of the average class consists of a discussion of concepts and homework conducted by the Graduate Assistant. Then new material is presented via TV using a pre-taped lecture which is transmitted from the TV studio over a closed-circuit system. Printed lecture notes are distributed at each lecture, allowing the students to concentrate fully on the TV presentation.

Extreme care went into the preparation of the TV lectures in an attempt to meet professional standards of TV production. Three faculty members of the mathematics department were assigned, as their sole teaching duties for $1\frac{1}{2}$ years, the making of a series of 63 video tapes. The Speech and Drama Department provided a professional director, a well-equipped studio, and a seven-man technical staff. On the average, the members of the mathematics department devoted about 75 man-hours to the making of each lecture. Although one person had the primary responsibility for each lecture, the lectures were thoroughly discussed by the other faculty members working on the course. Careful consideration was given to content, clarity, relevance, and style. Full use was made of the capabilities of the TV media by using slides, mats (subtitles), film clips, and imaginative props. As an example, part of a lecture on an application of mathematics to medicine was filmed at the National Institute of Health, with the commentary given by the director of a cancer research project.

Professor J. S. Mamelak of the Community College of Philadelphia, speaking also on the topic of handling large classes, suggested that modern display technics, closed circuit TV films, etc. encourage large class teaching. Such teaching tends to be well-organized and authoritative. Students learn factual information in large classes as well as they do in small classes. The principal drawbacks of teaching large classes are in the effective areas of student learning: lack of contiguous feedback, loss of interest and motivation. The traditional recitation method is frequently self-defeating. The expenditure on large class instruction is not fully offset by the use of assistants in recitation classes and the ability of young graduate students to administer to the broad demands of the student body is frequently questionable.

Some possibilities for using modern media technology to enhance teaching to large classes were indicated. The use of frequent testing as a motivator and objective evaluator of course content and methodology is possible with the current state of knowhow in all areas related to testing. The use of two-way radio and closed circuit TV technics to organize discussion periods in large classes under the direction of the principal lecturer to offset the weaknesses of the recitation arrangement is feasible. Finally, the increased use of numerical technics in the teaching of mathematics is suggested. Computer technology cannot be used in the feedback stage of learning today, but it can be used as a storage and retrieval device for the solution of assigned problems. A numeric program is largely self-checking; stress on numeric and algorithmic technics automatically changes the character of support required by the student in learning problem-solving.

MEETINGS OF OTHER ORGANIZATIONS

The American Mathematical Society held its sessions from Tuesday afternoon through Friday. There were two sets of Colloquium Lectures. Professor Lipman Bers of Columbia University gave one set, entitled "Uniformization, Moduli, and Kleinian

Groups," on Tuesday at 1:00 P.M., and on Wednesday, Thursday, and Friday at 8:45 A.M. Professor Armand Borel of the Institute for Advanced Study gave the other set, entitled "Algebraic Groups and Arithmetic Groups" on Tuesday at 2:15 P.M. and on Wednesday and Thursday at 10:00 A.M., and on Friday at 11:15 A.M.

The AMS Committee on Employment and Educational Policy presented a panel discussion on Tuesday afternoon at 3:30 P.M. with Professor W. L. Duren of the University of Virginia as moderator and Professors J. W. Jewett of Oklahoma State University and G. S. Young of the University of Rochester as panelists. Members of the panel discussed current problems of employment of Ph.D. mathematicians and sought the views of the mathematical public on these problems, prospects for the future, and consequences for graduate programs in mathematics.

Invited addresses were given by Professor J. M. Boardman of Johns Hopkins University on Friday at 10:00 A.M. on "Infinite Loop Spaces, Trees, and the Bar Construction," by Professor F. E. Browder of the University of Chicago on Thursday at 1:30 P.M. on "Nonlinear Functional Analysis," by Professor J. W. Robbin of the University of Wisconsin, Madison, on Friday at 1:30 P.M. on "Conjugacy Problems in Discrete Dynamical Systems," and by Professor Benjamin Weiss of the Hebrew University, Jerusalem, Israel, on Thursday at 2:45 P.M. on "Recent Progress on the Isomorphism Problem in Ergodic Theory."

The Pi Mu Epsilon Fraternity held sessions for contributed papers on Tuesday at 3:15 P.M. and on Wednesday at 10:40 A.M. in Room 312-313-314 of the Conference Center. A banquet was held on Tuesday at 6:30 P.M. in Dining Room B of the Hetzel Union Building. At this banquet, Professor J. W. Randolph of West Virginia University spoke on "Reflections on the Worth of Mathematics." A Dutch-treat breakfast meeting for Pi Mu Epsilon members was held on Wednesday at 8:00 A.M. in the Waring Hall Dining Room.

The Governing Council of Mu Alpha Theta, the national high school and junior college mathematics club, met on Wednesday at 9:00 A.M. in Room 113 of the Conference Center.

ARRANGEMENTS, ENTERTAINMENT, AND RECREATION

The Committee on Arrangements consisted of D. C. Rung, Chairman; H. L. Alder, Mrs. Patricia Axt, R. G. D. Ayoub, W. H. Gottschalk, T. J. Grilliot, Michael Keenan, S. F. Mack, T. D. Parsons, Mrs. Pilar Ribeiro, J. E. Schneider, G. L. Walker, Joseph Warren, S. W. Williams, T. J. Worgul.

Registration headquarters were located in the lobby of the Conference Center, the J. Orvis Keller Building. Dormitory rooms and cafeteria facilities were provided by Pennsylvania State University. Book exhibits and exhibits of educational media were displayed in Room 114 and 115 on the main floor of the Conference Center.

A chicken barbecue was held on Wednesday at 5:00 P.M. in the University Skating Pavillion. A beer party, sponsored by the Society for Industrial and Applied Mathematics, was held on Wednesday at 8:00 P.M. at Skimont Lodge. At both of these events, a local German band provided entertainment. A lecture on the Amish, entitled "Barndoor Britches and Shoo-fly Pie" was given by Dr. Maurice Mook on Tuesday at 7:30 P.M. Tours of the Amish market in Belleville were arranged for Wednesday, with two busses leaving at 8:45 A.M., one at 9:30 A.M., and an additional bus in the afternoon. A Chess Exhibition was held on Thursday at 7:30 P.M. in the Hetzel Union Building. Grand Master Donald Byrne of the English Department of Pennsylvania State University took two hours to defeat 14 opponents and draw with two others. The draws were with Bob Garrett, a freshman in mathematics at Pennsylvania State University, and Professor Emeritus Orrin Frink of Pennsylvania State University. Conducted tours of the University flower gardens were available.

HENRY L. ALDER, *Secretary*

MAY MEETING OF THE MISSOURI SECTION

The annual meeting of the Missouri Section of the MAA was held at Missouri Southern College, Joplin, on April 30 and May 1, 1971; seventy-five persons were in attendance.

Professor Charles Stuth, Section Vice-Chairman, presided at the Friday afternoon session, during which Professor A. B. Willcox gave the invited address, "England was Lost on the Playing Fields of Eton: A Parable for Mathematics," and the following papers were presented:

1. *On Schauder decompositions, two norm spaces and pseudo reflexivity*, by P. K. Subramanian, Missouri Southern College.
2. *The lattice of faces of a convex cone II*, by G. P. Barker, University of Missouri, Kansas City.
3. *A note on topology*, by Troy Hicks, University of Missouri, Rolla.
4. *A geometric introduction to stability theory and Liapunov functions*, by Stephen Bernfeld, University of Missouri, Columbia.
5. *Indefinite Finsler spaces*, by J. K. Beem, University of Missouri, Columbia.
6. *Criteria involved in the formulation of definitions involving sets*, by Henry Polowy, Lincoln University.

Professor Rochelle Boehning, Section Chairman, presided at the Saturday session, during which Professor J. W. Keesee gave the invited address, "Weakly Continuous Cohomology Theories." Also a panel discussion on Accreditation and Certification was presented by: Professor Glen Haddock, moderator, and panel members, Paul Burcham, University of Missouri, Columbia; L. T. Shiflett, Southwest Missouri State College; Ray Balbes, University of Missouri, St. Louis; and Charles Stuth, Stephens College.

At the business meeting, the following officers were elected for 1971–1972: Professor Charles Stuth, Stephens College, Chairman; Professor Fred Wilke, University of Missouri, St. Louis, Vice-Chairman; and Professor Troy Hicks, University of Missouri, Rolla, Secretary-Treasurer.

JACK JOLLY, *Secretary-Treasurer*

MAY MEETING OF THE ROCKY MOUNTAIN SECTION

Weber State College, Ogden, Utah, hosted the fifty-fourth Annual Meeting of the Rocky Mountain Section of the MAA on May 7 and 8, 1971. There were 65 registrants, including Professor W. N. Smith, of the University of Wyoming, the Sectional Governor, and Professor R. W. Irvine of Weber State College, the Section Chairman. The invited address, "Paths on Polyhedra," was delivered by Professor Victor Klee of the University of Washington, President of the Association. H. P. Hofmann, Academic Vice-President of Weber State College, welcomed the Section at the banquet Friday evening.

The following officers were elected at the business meeting: Chairman, C. A. Swanson, Southern Colorado State College; Vice-Chairman, Robert Gutzman, Colorado School of Mines; Second Vice-Chairman, C. N. Podraza, Northeastern Junior College; Secretary-Treasurer, D. J. Sterling, Colorado, College.

The following four papers were read at the invitation of the program committee:

1. *Recent developments in geometric topology*, by L. C. Glaser, University of Utah.
2. *Calculus: CUPM's unused version*, by Ben Roth, University of Wyoming.
3. *Accreditation and certification*, by D. J. Sterling, Colorado College.
4. *Computer graphics and the head-mounted display*, by D. L. Vickers, University of Utah.

Ten papers were contributed and read on the program:

1. *A relation between π and greatest common divisors*, by David Ballew, South Dakota School of Mines and Technology.
2. *A generalization of a conjecture of Erdös*, by R. B. Crittendon, Portland State University.
3. *A generalized Riemann-Stieltjes integral*, by M. L. Klasi, South Dakota School of Mines and Technology.
4. *Self-directed study in mathematics*, by K. F. Klopfenstein*, Wilson Brumley, Darrell Perkins, Colorado State University.
5. *Partitions of a matrix*, by A. D. Porter, University of Wyoming.
6. *Generalized inverses of group homomorphisms*, by D. W. Robinson, Brigham Young University.
7. *Categorical methods applied to Pontryagin duality*, by D. W. Roeder, Colorado College.
8. *Incidence algebra and GF* $[q, x]$, by L. E. Shader, University of Wyoming.
9. *On an existence theorem for boundary value problems*, by W. G. Sutton*, South Dakota School of Mines and Technology, J. H. George, University of Wyoming.
10. *Arc length and the mean value theorem*, by S. G. Wayment, Utah State University.

In addition to the above papers, an exhibit of textbooks for use in the junior and community college curriculum was presented with the generous help of the following publishers: Addison-Wesley; Harcourt Brace Jovanovich; Harper and Row; Prindle, Weber and Schmidt; Scott Foresman; and Van Nostrand Reinhold.

D. J. STERLING, *Secretary-Treasurer*

CUPM AND THE MATHEMATICAL SOCIAL SCIENCE BOARD

The Mathematical Social Science Board and the Committee on the Undergraduate Program in Mathematics are seeking interesting problems or illustrative examples, from each of the social sciences, whose solutions and study make use of ideas and techniques from one or more of the following topics in undergraduate mathematics: sets and relations, differential and integral calculus, matrices and linear algebra, and probability.

We propose to collect such examples into a book mainly to be used by mathematics teachers and students as a source (1) of current social science applications of mathematics and (2) of material for textbook and classroom exercises to illustrate how topics in collegiate mathematics arise in a social science context. We also plan to include annotated bibliographies of articles and books involving applications of mathematics to the various social sciences.

The most preferred contribution would be an exposition giving (a) the social science problem and its background, (b) the reduction of the problem to mathematical form, (c) the mathematical analysis, perhaps with associated numerical results obtained on a computer, and (d) the meaning and insights provided by the mathematical analysis when related back to the original social science problem. Less desirable, but still very welcome, would be a reprint including material from which such an exposition could be extracted. References to the literature would also be helpful.

The CUPM-MSSB Project Committee presently consists of the following persons: D. W. Bushaw, Samuel Goldberg, Harold Kuhn, R. D. Luce, Henry Pollak.

Please send contributions to: CUPM-MSSB Project, Post Office Box 1024, Berkeley, California 94701.

ACADEMIC MEMBERS ELECTED INTO THE ASSOCIATION

In accordance with the amendment adopted at the business meeting of the Association at Stillwater on August 30, 1961, the Board of Governors at its meeting at Pennsylvania State University, University Park, Pennsylvania, on August 29, 1971, elected to membership the twentieth set of applicants for academic membership (for election of the other nineteen sets, see the March and December issues of 1969, the April and November issues of 1970, and the April issue of 1971). Approval for election was given to the following seven applicants for academic membership:

Butte College, Durham, California
Delaware Valley College of Science and Agriculture, Doylestown, Pennsylvania
Federal City College, Washington, D. C.
Greater Hartford Community College, Hartford, Connecticut
Mankato State College, Mankato, Minnesota
Marion College, Fond du Lac, Wisconsin
Queen's University, Kingston, Ontario, Canada

HENRY L. ALDER, *Secretary*

Page 1

U S POSTAL SERVICE
STATEMENT OF OWNERSHIP, MANAGEMENT AND CIRCULATION
(Act of August 12, 1970 Section 3685 Title 39, United States Code)

SEE INSTRUCTIONS ON PAGE 2 (REVERSE)

1. TITLE OF PUBLICATION
THE AMERICAN MATHEMATICAL MONTHLY

2. DATE OF FILING
October 1, 1971

3. FREQUENCY OF ISSUE
Monthly- except July and August

4. LOCATION OF KNOWN OFFICE OF PUBLICATION *(Street, city, county, state, ZIP code) (Not printers)*
1225 Connecticut Avenue, N. W., Washington, D. C. 20036

5. LOCATION OF THE HEADQUARTERS OR GENERAL BUSINESS OFFICES OF THE PUBLISHERS *(Not printers)*
Same as above

6. NAMES AND ADDRESSES OF PUBLISHER, EDITOR, AND MANAGING EDITOR

PUBLISHER *(Name and address)*
Mathematical Association of America, 1225 Connecticut Avenue, N.W.,Washington,D.C.

EDITOR *(Name and address)*
Harley Flanders, Dept. of Math., Tel Aviv University, Ramat-Aviv, ISRAEL

MANAGER EDITOR *(Name and address)*
Raoul Hailpern, Math Ass'n of America, SUNY at Buffalo, Buffalo, NY 14214

7. OWNER *(If owned by a corporation, its name and address must be stated and also immediately thereunder the names and addresses of stockholders owning or holding 1 percent or more of total amount of stock If not owned by a corporation, the names and addresses of the individual owners must be given If owned by a partnership or other unincorporated firm, its name and address, as well as that of each individual must be given)*

NAME	ADDRESS
Mathematical Association of America	1225 Connecticut Avenue, N.W.Washington,D.C. 20036

8. KNOWN BONDHOLDERS, MORTGAGEES, AND OTHER SECURITY HOLDERS OWNING OR HOLDING 1 PERCENT OR MORE OF TOTAL AMOUNT OF BONDS, MORTGAGES OR OTHER SECURITIES *(If there are none, so state)*

NAME	ADDRESS
NONE	

9. FOR OPTIONAL COMPLETION BY PUBLISHERS MAILING AT THE REGULAR RATES (Section 132.121, Postal Service Manual)

39 U. S. C. 3626 provides in pertinent part "No person who would have been entitled to mail matter under former section 4359 of this title shall mail such matter at the rates provided under this subsection unless he files annually with the Postal Service a written request for permission to mail matter at such rates."

In accordance with the provisions of this statute, I hereby request permission to mail the publication named in item 1 at the reduced postage rates presently authorized by 39 U S. C. 3626.

(Signature and title of editor, publisher, business manager, or owner)

10. FOR COMPLETION BY NONPROFIT ORGANIZATIONS AUTHORIZED TO MAIL AT SPECIAL RATES *(Section 132.122, Postal Manual)*
(Check one)

The purpose, function, and nonprofit status of this organization and the exempt status for Federal income tax purposes

☒ Have not changed during preceding 12 months

☐ Have changed during preceding 12 months

(If changed, publisher must submit explanation of change with this statement.)

11. EXTENT AND NATURE OF CIRCULATION	AVERAGE NO. COPIES EACH ISSUE DURING PRECEDING 12 MONTHS	ACTUAL NUMBER OF COPIES OF SINGLE ISSUE PUBLISHED NEAREST TO FILING DATE
A. TOTAL NO. COPIES PRINTED *(Net Press Run)*	23,420	23,250
B. PAID CIRCULATION 1. SALES THROUGH DEALERS AND CARRIERS, STREET VENDORS AND COUNTER SALES	NONE	NONE
2. MAIL SUBSCRIPTIONS		
C. TOTAL PAID CIRCULATION	21,642	21,078
D. FREE DISTRIBUTION BY MAIL, CARRIER OR OTHER MEANS 1. SAMPLES, COMPLIMENTARY, AND OTHER FREE COPIES	50	50
2. COPIES DISTRIBUTED TO NEWS AGENTS, BUT NOT SOLD		
E. TOTAL DISTRIBUTION *(Sum of C and D)*	21,692	21,128
F. OFFICE USE, LEFT-OVER, UNACCOUNTED, SPOILED AFTER PRINTING	1,728	2,122
G. TOTAL *(Sum of E & F—should equal net press run shown in A)*	23,420	23,250

I certify that the statements made by me above are correct and complete.

(Signature of editor, publisher, business manager, or owner)
A. B. Willcox, Executive Director

PS Form 3526 July 1971

REPORT OF THE TREASURER FOR THE YEAR 1970

Included here is a summary of the report of the Treasurer of the Association for the year 1970. The report has been approved by the Finance Committee and accepted by vote of the Board of Governors. Any member of the Association who wishes to have a copy of the full report may obtain one by writing to the Washington office of M.A.A.

In contrast with the two previous years, in which we operated with deficits, this year the report shows a small excess of income over expenditures of $3,114.15.

In conformity with current accounting principles, securities are listed at cost rather than market value.

ASSETS	*Dec. 31, 1969*	*Dec. 31, 1970*
Cash	$ 21,935.76	$ 86,124.57
Securities	155,005.80	155,629.00
Accounts Receivable	31,206.58	49,908.64
Furniture and Equipment	17,061.17	17,586.14
Prepaid Expense	0	3,017.00
Total Assets	$225,209.31	$312,265.35
LIABILITIES		
Accounts Payable	19,178.44	24,595.14
Unearned Income		
Dues	117,984.35	169,319.79
Subscriptions	36,013.55	49,347.18
Other	0	6,705.00
NSF Fund	12,128.32	18,059.83
High School Contest Fund	(5,618.73)	(4,399.12)
Total Liabilities	$179,685.93	$263,627.82
Assets minus Liabilities	$ 45,523.38	$ 48,637.53

INCOME	
Dues and Initiation Fees	$201,336.28
Subscriptions, sale of publicationsand advertising	138,809.94
Dividends and Interest	7,782.64
Contributions	6,127.80
Registration Fees—National Meetings	3,953.62
Indirect Costs (NSF)	79,587.15
Other	1,626.25
Total Income	$439,223.68
EXPENDITURES	
Salaries	169,981.08
Office Expenses	57,159.14
Publications	142,842.41
Travel	23,816.81
Taxes	4,872.89
Dues and Contributions	13,770.00
Joint ventures	14,072.93
Awards and Grants	4,837.31
Other	4,756.96
Total Expenditures	$436,109.53
Income over Expenditures	$ 3,114.15

E. A. CAMERON, *Treasurer*

CALENDAR OF FUTURE MEETINGS

Fifty-fifth Annual Meeting, Las Vegas, Nevada, January 19–21, 1972.

Fifty-third Summer Meeting, Dartmouth College, Hanover, New Hampshire, August 28–30, 1972.

The following is a list of the Sections of the Association with dates of future meetings so far as they have been reported to the Editorial Director.

ALLEGHENY MOUNTAIN, Pennsylvania State University, Altoona, May 5–6, 1972.

FLORIDA, Central Florida Junior College, Ocala, March 17–18, 1972.

ILLINOIS, Lake Forest College, May 12–13, 1972.

INDIANA

IOWA, University of Iowa, Iowa City, April 28, 1972.

KANSAS, Washburn University, Topeka, March 24–25, 1972.

KENTUCKY, Georgetown University, Georgetown, Spring 1972.

LOUISIANA-MISSISSIPPI, Millsaps College, Jackson, Mississippi, February 18–19, 1972.

MARYLAND-DISTRICT OF COLUMBIA-VIRGINIA

METROPOLITAN NEW YORK

MICHIGAN, Oakland University, Rochester, May 5–6, 1972.

MISSOURI, Stephens College, Columbia, May 5–6, 1972.

NEBRASKA, University of Nebraska at Omaha, Omaha, April 21–22, 1972.

NEW JERSEY

NORTH CENTRAL

NORTHEASTERN

NORTHERN CALIFORNIA, California State College at Hayward, Hayward, February 5 1972.

OHIO, Wittenberg University, Springfield, April 28–29, 1972.

OKLAHOMA-ARKANSAS, State College of Arkansas, Conway, Arkansas, March 10–11, 1972.

PACIFIC NORTHWEST, University of Washington, Seattle, Washington, June 16–17, 1972.

PHILADELPHIA

ROCKY MOUNTAIN, Southern Colorado State College, Pueblo, May 5–6, 1972.

SOUTHEASTERN, Samford University, Birmingham, Alabama, March 24–25, 1972.

SOUTHERN CALIFORNIA, California Institute of Technology, Pasadena, March 11, 1972.

SOUTHWESTERN, University of New Mexico, Albuquerque, Spring 1972.

TEXAS, Southwest Texas State University, San Marcos, April 1972.

UPPER NEW YORK STATE

WISCONSIN, Wisconsin State University, Stevens Point, April 28–29, 1972.

FUTURE MEETINGS OF OTHER ORGANIZATIONS

AMERICAN ASSOCIATION FOR THE ADVANCEMENT OF SCIENCE, Philadelphia, December 26–31, 1971.

AMERICAN MATHEMATICAL SOCIETY, Las Vegas, Nevada, January 17–20, 1972.

AMERICAN SOCIETY FOR ENGINEERING EDUCATION

ASSOCIATION FOR COMPUTING MACHINERY, Boston, Massachusetts, August 14–16, 1972.

ASSOCIATION FOR SYMBOLIC LOGIC, Statler Hilton Hotel, New York City, December 27–28, 1971.

CENTRAL ASSOCIATION OF SCIENCE AND MATHEMATICS TEACHERS, Chicago, Illinois, November 16–18, 1972.

FIBONACCI ASSOCIATION

INSTITUTE OF MATHEMATICAL STATISTICS

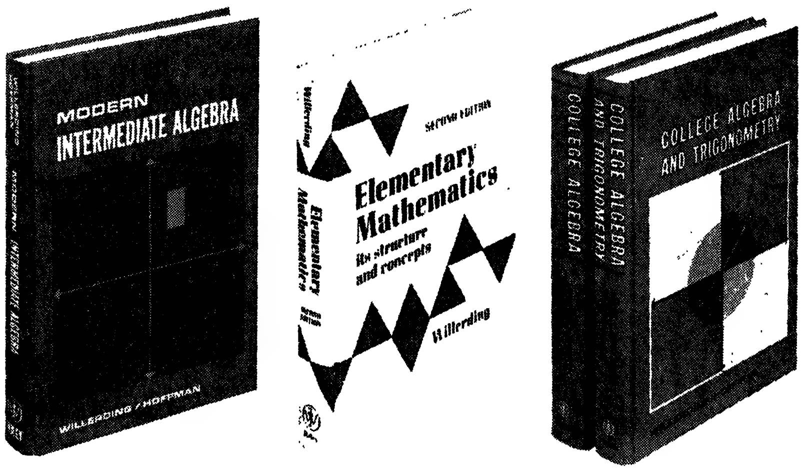
MU ALPHA THETA

NATIONAL COUNCIL OF TEACHERS OF MATHEMATICS, Chicago, Illinois, April 16–19, 1972.

OPERATIONS RESEARCH SOCIETY OF AMERICA, Jung Hotel, New Orleans, April 26–28, 1972.

PI MU EPSILON, Dartmouth College, Hanover, New Hampshire, August 29–30, 1972.

SOCIETY FOR INDUSTRIAL AND APPLIED MATHEMATICS, Benjamin Franklin Hotel, Philadelphia, June 12–14, 1972 (20th Anniversary Celebration).


MODERN
INTERMEDIATE ALGEBRA
WILLERDING / HOFFMAN
Elementary Mathematics
Its structure and concepts
Willerding
COLLEGE ALGEBRA AND TRIGONOMETRY
COLLEGE ALGEBRA

GEORGE BANTA COMPANY, INC., MENASHA, WISCONSIN, U.S.A.

INDEX TO VOLUME 78, 1971

THE AMERICAN MATHEMATICAL MONTHLY

AUTHOR INDEX

KEY WORDS AND PHRASES INDEX

PROBLEMS AND SOLUTIONS

PROBLEMS PROPOSED

PROBLEMS SOLVED

SOLUTIONS

Numbers in **boldface** type refer to problems, those in lightface, to pages.

REVIEWS

Abbreviations: (TR)—Telegraphic Review; (NP)—Notable Paper.
Names of authors are in ordinary type, those of reviewers in capitals.

NEWS AND NOTICES

PERSONAL ITEMS

GENERAL INFORMATION

NECROLOGY

REPORTS AND ANNOUNCEMENTS OF THE ASSOCIATION AND ITS SECTIONS

MEETINGS AND ANNOUNCEMENTS OF THE ASSOCIATION

MEETINGS OF ITS SECTIONS

THE AMERICAN MATHEMATICAL MONTHLY

(FOUNDED IN 1894 BY BENJAMIN F. FINKEL)

THE OFFICIAL JOURNAL OF

THE MATHEMATICAL ASSOCIATION OF AMERICA, INC.

VOLUME 78 NUMBER 10

CONTENTS

(Continued on inside cover)

DECEMBER 1971

NOTICE TO AUTHORS

Specialized research is usually unsuitable; see *Statement of Policy* (vol. 76, p. 2). Manuscript preparation: Please use the *Manual for Monthly Authors* (vol. 78, p. 1) and follow the format in current issues of the Monthly. Manuscripts should be typewritten, triple-spaced with wide margins; submit two copies and keep one for protection against loss.

Backlog: Main Articles 7 months, Math. Notes 8 months, Research Problems 6 months, Classroom Notes 7 months, Math. Education 6 months.

EDITORIAL CORRESPONDENCE AND MAIN ARTICLES: to Harley Flanders, American Mathematical Monthly, Tel Aviv University, Ramat Aviv, Israel (see Notice, vol. 77, 1970, p. 555); NOTES, etc.: to the corresponding Associate Editor; ADVERTISING CORRESPONDENCE: to Raoul Hailpern, Mathematical Association of America, SUNY at Buffalo, Buffalo, N. Y. 14214; CHANGE OF ADDRESS and SUBSCRIPTIONS: to A. B. Willcox, Mathematical Association of America, 1225 Connecticut Ave., N.W., Washington, D. C. 20036.

Annual dues for members of the Association (including a subscription to the American Mathematical Monthly) are $12.50. For nonmembers the subscription price is $18.00.

Published by the Association at Washington, D. C., and Menasha, Wisconsin during the months of January, February, March, April, May, June–July, August–September, October, November, December.

Second-class postage paid at Washington, D. C., and additional mailing offices. Acceptance for mailing at special rate of postage provided for in the Act of February 28, 1925, embodied in Paragraph 4, Section 538, P. L. and R., authorized April 1, 1926.

PRINTED IN THE UNITED STATES OF AMERICA

1871: OUR STATE OF MATHEMATICAL IGNORANCE

H. B. GRIFFITHS, The University of Southampton (England)

1. What does the title mean? The title of this lecture strikes me as highly enigmatic, and I accepted it from the Committee because I was curious to investigate what it meant. (1) Surely, some comparison between 'then' and 'now' was wanted, and a quick but superficial response is to quote the information [**15**] that in 1868, some 800 titles of mathematics publications are listed, as against 13,000 for 1965. 'Hence,' as the bright brash adolescents say, 'QED, $\cdots$, trivial,' and we may as well spend the rest of the time discussing the unsolved problems of mathematics. Many of us, however, would find such an answer unsatisfactory, because mathematics is for most people a social activity. Hence, we shall understand it and ourselves better, especially in relation to our teaching, if we think of mathematics as an organism, not existing *in vacuo*, but as something generated by people of a common craft, that of mathematician. Also, what is mathematical ignorance? Lack of knowledge in one direction may be compensated by a surfeit in another. Thus, I want to approach the problem from the point of view of the discipline of Mathematical Education, rather than that of Mathematics alone. I take it that the Committee wanted me to compare the mathematical ignorance of the 'average chap' in 1871 with that of the 'average chap' of today, using the term to include females and also those who may not perhaps be members of this Association. The trouble with this project is to describe the 'average chap' of 1871, and I can only hope to do it inadequately in the time available; it would take several lectures, and a lot more reading than I have so far done to do justice to this difficult and fascinating problem. But to attempt the description will be helpful to understand our average chap of to-day, and why his mathematics is what it is. Before I go into that, however, let us look at the mathematics of 1871. For brevity I must grossly compress what must be known to many of you through the works of Bell [**4, 5**].

2. Mathematical trends in 1871. Within the trade, commercial arithmetic, algebra, geometry in the style of Euclid, and Newtonian Calculus were fairly common knowledge in 1871. Among the relatively few creative mathematicians,

---

This article is based on a lecture presented at the Centenary Conference of The Mathematical Association in London, April, 1971. (At the Conference, the representative of the MAA presented a certificate and announced the agreement to a reciprocity agreement between our MAA and the British MA.)

Professor Griffiths did his Ph.D. work at the University of Manchester under M. H. A. Newman. He has held positions at the Universities of Aberdeen, Bristol, Birmingham, and presently Southampton, where he is Chairman of the Mathematics Department. He has spent leaves at the Institute for Advanced Study and the Courant Institute. In addition to his research interests in topology and function theory, he has been active in mathematical education. He is the author of *Classical Mathematics* (with Peter Hilton, Van Nostrand-Reinhold, 1970) and *A Comparison to A-Level Mathematics* (with Vernon Armitage, SMP Series, Cambridge University Press, 1969). *Editor.*

(1) Numbers in parentheses refer to the notes at the end of article.

additional pregnant notions were evolving, whose importance is much easier for us to see, looking back. First, there was Mathematical Physics. In 1828 Green had published his essay *The Application of Mathematical Analysis to the Theories of Electricity and Magnetism.* This began modern mathematical physics in Britain, and by 1871 it was in the hands of men like Maxwell, Kelvin, Tait, Rayleigh, Helmholtz, Kirchhoff, Gibbs (in USA) and, by the 1890's, Poincaré. Much of the Physics had later to be revised in the light of Relativity and Quantum Physics, but the associated mathematical ideas are still with us, very much alive. Briefly, the attempts to understand Nature led to descriptions of it that we now call *mathematical models*; and an important feature of those models was the occurrence of differential and integral equations. These we are still learning to solve. Questions of Probability and Statistics were in the air too, because Statistical Thermodynamics and Actuarial Mathematics were beginning, with Mathematical Genetics not far behind (Mendel's paper was published in 1865). To get the tools for solving the equations and constructing the models, we must turn to Pure Mathematics. By 1871, the craft had left a rich legacy of finished mathematics, together with some half-emergent ideas that seem to us now of even greater importance. Let me enumerate them:

(A) *The notion of algorithm.* Gauss, Wantzel, Abel, Galois, and others had discussed such questions as the possibility or otherwise of constructing (with ruler and compass) the regular polygons, of trisecting angles and duplicating the cube, and the solution of equations by radicals. The notion of a 'procedure' or 'algorithm' for carrying out the proposed constructions had to be analysed. Taken with the work of Babbage (1792–1871) on his 'analytical engine', together with the later, descendent ideas of the Logicians of the 1930's, these notions were the right ones for developing the electronic digital computer, with all its potential effects on contemporary life.

(B) *The notion of algebraic structure.* Gauss and Galois, especially, had had to develop the notions of *field* and *group* to do their work. Although some German Universities had lecture courses on Groups from the 1850's, Jordan's book *Traité des Substitutions* was the first to publicise the notions widely: it appeared in 1871. Quite independently, Hamilton in 1852 developed his algebra of Quaternions to apply to Physics, while Boole in 1854 published his *Laws of Thought* in which he had developed the algebra of sets—Boolean algebra. From the complex numbers and the quaternions it was a natural step to investigate hypercomplex numbers, and by 1870 enough was known for Benjamin Pierce of Harvard to publish his book *Linear Associative Algebras.* Cayley and Sylvester, too, had investigated matrices and their algebraic structure, partly with a view to developing a language for geometry, which I will mention below.

(C) *The notion of function.* The theory of Fourier series had led Dirichlet in the 1830's to formulate a satisfactory notion of function which was more or less the modern one. It was a basic tool for the development of complex variable theory, and above all, for Riemann's Theory of functions (Riemann died in

1866). Here, one important idea was that of considering, not just special functions like sin x, cn x, $J_0(x)$, $\Gamma(x)$, in the 18th Century spirit; rather, one should consider the set of *all* (differentiable) functions from one domain of complex numbers to another, and one should ask for the 'typical', 'generic' structure of the functions of the entire set. This is a broadening of our horizons comparable to asking for a study of Man, rather than 'my neighbour'. Even more, Riemann replaced the domains and ranges of his families of functions by his 'Riemann surfaces': he had taken the first step from the flat plane(2) of complex numbers, to the 'curved' n-dimensional manifolds which Poincaré later found to be the proper models for studying differential equations.

(D) *The notion of geometric structure.* In the early 19th Century, Projective Geometry had flourished with Monge and Poncelet. Grassmann had invented his calculus of linear subspaces of n-dimensional Euclidean space, and algebra had to be developed for describing these rich new structures. From a philosophical point of view, however, the really liberating event was the discovery of non-Euclidean geometry by Gauss, Bolyai and Lobatchewsky: it was at last realised—though Gauss lacked the courage to face the publicity—that Euclidean geometry was but one possible model by which to describe the external world. L. Bers has given an illuminating idea(3) of this philosophical difference by comparing the 'Euclidean' attitude of the Declaration of Independence ('We hold these truths to be self-evident · · · ') with the 'non-Euclidean' attitude of Lincoln's Gettysburg Address of 1863 (' · · · a nation dedicated to the proposition that all men are created equal'). Perhaps the climate of the time was ripe for such ideas to generate in different places, but in Mathematics, the first to see the implications of this relativistic point of view was Riemann, who seized on Gauss's differential geometry of surfaces to incorporate it into his n-manifolds, with a 'Riemannian metric'. He even hinted that the proper model of force should be curvature, as Einstein later taught us. And by 1872, Klein had unified all these geometries within his Erlanger Programm, wherein he stressed that the important mathematical feature of any geometry is its associated group of symmetries and their invariants—here is the notion of function again. A wholly natural generalisation of Klein's ideas is the contemporary theory of categories and functors, and his message has been amply justified: *to study an object, observe how it maps into known objects and how they map into it.*

(E) *The arithmetisation of analysis.* By mid-century, Weierstrass, Heine, and others had invented the language of rigorous analysis. The older vague ideas of 'moving to a limit', and of 'infinitesimal', had made it difficult to give convincing proofs of the rules required for doing the increasingly more complicated calculus. Simple inequalities, the fearsome 'epsilonology' of contemporary undergraduates, got rid of the vagueness, and made good proofs possible. The cost was a certain alienation of the 'average chap' because more effort is needed to learn this language than to memorise the old murky proofs. Once the language was invented, however, it at once showed gaps in our understanding of the number system; and Dedekind and Cantor filled them in. Cantor was well

into his Theory of Sets by 1870, and he provided us with a serviceable working language which nevertheless was not precise enough to avoid the famous contradictions concerning the Infinite. These led to the later work of Peano (1858–1932), Bertrand Russell (born in 1872), A. N. Whitehead (1861–1947), K. Goedel, and the later logicians.

3. Developments in mathematics. In the succeeding century, these ideas have been digested and extended, to have great potential effects on our contemporary average chap. Firstly, the improvement in precision of language made it possible (granted some ingenuity also) to prove many suspected results, like the transcendence of π and e, and the Prime Number Theorem; the Dirichlet Problem in Potential Theory could be dealt with properly, whereas Riemann had had to assume it; and such facts as the denumerability of the algebraic numbers and the non-denumerability of the reals could be established. At last, too, the vital notion of *isomorphism* could be expressed objectively.

Secondly, the language, and the multiplicity of structures, led to the *axiomatic method*, used first by Hilbert in his book (1899) 'The Foundations of Geometry' to put right the blemishes in Euclid that the new precise language had shown to exist. It was followed by the dictum of E. H. Moore, Professor in Chicago (1910) that when two theories possessed similar theorems, it was a mathematician's duty to uncover the common underlying structure (which is rather different from looking for an isomorphism between, say, groups; one looks for an isomorphism between languages). Thus, following the classical success of Potential Theory—the common structure of several theories of Mathematical Physics such as Electrostatics, Ideal fluids, $\cdots$,—more *applications* of the common abstract theory may be possible. This is the practical justification of the 'abstract' approach. The resulting style of doing mathematics got its greatest exponent in the work of N. Bourbaki, whose enormous and still growing book is effecting a great unification of mathematics by its means. His technique has for twenty years now been copied by younger mathematicians, to solve many problems, that would have been too complicated for classical methods. In particular, the axiomatic method helps to separate the details of a complicated construction of a mathematical tool, from the use to which the tool is put. For an example, see the proof of Brouwer's Fixed point theorem in [11], p. 441. The method also forces a revision in our attitude to what an acceptable answer to a question should be: Dedekind's answer to the question 'What is a number?' was to formulate a set of axioms that any good *system* of numbers should satisfy, demonstrate the existence of such a system, and leave it at that. And an engineer's emphasis creeps in, because of the question 'Can your axioms and hypotheses be verified in a finite number of steps—*are they computable*?'. Classification theorems arise, to say how many things satisfy a given set of axioms. All this is a far cry from the widely held contemporary misconception that axioms are simply things you vary at will: one concentrates on *significant* sets of axioms, significant because they describe abstractly some interesting concrete situation.

Thirdly, the language and the structures were available for the creation after 1900 of the branch of geometry called Algebraic Topology and its many offshoots. Following ideas of Euler, Listing and Tait (who invented Knot Theory), Poincaré had invented it, to deal with problems of celestial mechanics, when Riemann's two-dimensional methods became inadequate; but it took half a century of the efforts of many mathematicians, using the language of Set Theory, Logic, Analysis and Algebra, to get it sufficiently organised and flexible for a strong attack to be made on the problems envisaged by Poincaré. The resulting weapon of attack is called 'Global Analysis': for an example of the influence of this approach in Mechanics, see the book by Abraham [**1**]. It is also being used by Thom and others to extend the work of D'Arcy Thompson [**21**] in mathematical biology with some potentially profound consequences. Related, though as yet more elementary, structures are being used in Economics, Theory of Games, and Operational Research. Certainly, the Applied Mathematicians now have been provided with a fantastic range of structures from which to model aspects of Nature.

Hardly anyone understood this potential of the axiomatic method in 1871, of course, although in England, A. H. Clifford [**9**] had understood some of the implications of non-Euclidean geometry and tried to lay this forth in his popular writings. He certainly appreciated the implications for the relativity of knowledge in the sense that one should speak, not of absolute laws of behaviour in nature, but only of behaviour which approximated to an appropriately chosen mathematical model.

4. Mathematical education. Coming to the related discipline of Mathematical Education, (which concerns the selection and communication of mathematical topics subject to various constraints) all this activity has made available for us various languages and models.

The language enabled adult mathematicians to talk to each other more efficiently about basic mathematical notions such as Number, and therefore it would seem reasonable to use some version of that language for explaining difficulties to children. Sometimes it may be more efficient to talk about a *model* of a system to impart understanding of that system, as with Cuisenaire rods and the system of rationals. To choose the language or model requires hard thought followed by experimental observation of pupils; hard criticism is necessary too, not the off-the-cuff prejudices of certain supporters of Dr. Hammersley [**12**] and the Black[(4)] Papers [**6**], who have rarely thought as *seriously* about mathematical education as they may have about mathematics. Since language is designed to model previously incoherent, vague feelings, it is perhaps not surprising that ignorant hostility should even claim the 'Silent Majority' as allies (as did a recent *Gazette* correspondent (4a)); but one reason for including Mathematics in a General education is exactly to tame those visceral feelings of murder and obscenity that the 'Silent Majority' is alleged by its political spokesmen to stand for, when they want to justify any wickedness.

Of the five principal notions described above, today's average chap is affected only by the third, the notion of a function, but only in the restricted eighteenth century sense of 'an expression'; and in spite of the notice taken of Klein's book [**13**] *Elementary Mathematics from an Advanced Standpoint*, wherein he hectors his audience of 19th century schoolmasters and lays great stress on the notion of function. True, University graduates have met the fifth notion, the arithmetisation of analysis, but it has had little effect on their later teaching in schools and many say quite frankly that it left them cold. Something of the other notions can be found in the newer mathematics projects, but as yet these leave many average chaps untouched, and in any case the ideas may be received in a degenerate form. The pessimism of this last remark derives from my recent reading of past enthusiasts, that I shall describe in a moment. Now, these enthusiasts achieved some things, but failed in the crucial matter of making mathematical activity a creative activity for the average chap. For, even today, the only mathematical activity known to our average chap (outside Primary School!) is that of reproducing work for examinations. He does not need complex language to express fine distinctions and formulate questions, since the only mathematical questions he meets are those formulated for him by examiners who already know an answer. This is clearly a poor way to grasp pregnant notions, since he never really needs to deploy them. Before we can hope to improve on what we do, we must ask how our present practices came about. And this brings me back to trying to describe our average chap and the trends that made him what he is.

5. General climate of thought in 1871. So what was it like to be an average chap in 1871? For brevity and simplicity I shall have to assume that he was British.

Britain was a much less homogeneous country then than now, both through its class divisions and those of nationality. For example, a good approximate picture of the 'average chap' in Scotland can be got from Knott's *Life and Scientific Work of P. G. Tait* [**14**], although he writes about the very *unaverage* trio of friends, Tait, Kelvin, Maxwell, and their many acquaintances—Stokes, Helmholtz, Huxley, and others. That was a time when one of Tait's assistants[(5)] could become Professor of Hebrew in Aberdeen, be dismissed in 1881 for unorthodox views, and later become Professor of Arabic in Cambridge. The integrity of men like Tait is presumably partly responsible for the academic freedom we now enjoy, but it is curious to see how the Scots for social reasons looked for educational leadership to England, (and sent their leaders' sons here for education). Brevity makes it necessary for me to concentrate now upon the English situation, with its quite different educational tradition.

Our English average chap would, therefore, think of himself in 1871 as being at the centre of a Universe, presided over by Queen Victoria and the Prime Minister Gladstone, with considerable aid from many able servants of the Imperial Ideal, including God as honorary citizen of the British Empire. Assuming that he read his Newspapers he would be worried by the proclamation of

the German Empire at the conclusion of the Franco-Prussian War; and about the threat of competition from the Germans in trade. The Americans, too, were becoming a commercial threat even as they were recovering from the ravages of the Civil War which had ended six years previously. He could read of the political upheavals of the Paris Commune (whose effects are still with us studentwise) and possibly of the exploits of a General Bourbaki whose training in Africa had left him pathetically ill-equipped to deal with the new problems of the war with Prussia. In Japan, feudalism had just ended with the disbandment of the Samurai; and Westerners (including Knott) were being brought in, to promote the very civilised avowed intentions of the State. By contrast, the great powers of the West were still carving up Africa and behaving as badly in China.

Tolstoy had just completed *War and Peace*, Marx was in full spate and Lenin was just born, Dickens had just died, Mark Twain was 36, and Bernard Shaw was an adolescent. A whole galaxy of painters and romantic composers were alive but largely unknown to our average chap, because Thomas Edison, working on the electric lamp, had not yet developed the techniques of communication that would make their existence well-known. There was intellectual turmoil in Britain because of the effect of Darwin on religion, but the human mind seemed capable of infinite progress, and the certain dogmas of the scientists were confronted with equally naïve certainty by the counter-dogma of the church; perhaps for that reason Papal infallibility had just been proclaimed. On both sides, things always turned out less simple than was expected, as the Pope is now finding out and the scientists and mathematicians found out much longer ago. Religious Tests at Oxbridge had just been abolished, and these two Universities needed no longer to drive Dissenters to Durham, London, or Manchester; Sylvester could soon return to a Chair from Baltimore.

Illness was still risky, in 1871, but Pasteur was in his prime and Lister was beginning[6] his work with antiseptic surgery. Public Sanitation Acts were beginning to be passed and Trades Unions in Britain were legalised. The Forster Act of 1870 had made elementary education available for all, but we need only look at what led up to that Act, and the effect it has had upon us all here, to see how very difficult it is to decide on what was an average chap of 1871—much more difficult than it would be to delineate such a person today. It is essential—he and his administrators would insist—that we place him in one of three classes, that I shall call without definition 'the Leisured,' 'PB' (for Petit bourgeois) and 'the Working class.' Novelists indicate that he would prefer much finer graduations, but these three will suffice for now. An education system appropriate to each class was developed as the Nineteenth century wore on, but largely *planned by members of the Leisured class;* and each had its own attitude to mathematics. Briefly, Oxford and Cambridge with the Public Schools were for the Leisured class, private and grammar schools for the PB, and the Elementary schools for the Working class. (For simplicity I omit the other English Universities of the period.) Thus, each class might have contained its own 'average chap,' but each such would have aspirations and characteristics quite different from the others.

The history of British education since then is partly a history of the blurring of those distinctions. I can only give a crude simplification here: to fill out the picture I found the books of Armytage [**2**], Montgomery [**17**], and Wardle [**23**] most helpful, as well as the historical introduction to the Spens report [**19**] of 1938. But mathematics in particular has turned out to be an instrument of use in that blurring process.

6. The role of mathematics: examinations. Mathematics played its part in the blurring, firstly because knowledge of mathematics was compulsory in the newly formed competitive degrees of Oxford and Cambridge which had been established around 1800. Recall that these Universities had up to then fallen into a state of advanced decay as had the Public Schools and Grammar Schools; they began to stir again in the first third of the century. An interesting description of the decline of Mathematics in Cambridge, following the death of Newton, is given in Chapters V and VI of Rouse Ball [**3**]. He includes (p. 111) a revealing letter by Sir F. Pollock (Senior Wrangler in 1806) to De Morgan, in which Pollock claims he was the last 'fluxional' Senior Wrangler (though De Morgan recalled others). His letter concludes:

"My experience has led me to doubt the value of competitive examination. I believe the most valuable qualities for practical life cannot be got at by any examination—such as steadiness and perseverance. It may be well to make an examination part of the mode of judging of a man's fitness; but to put him into an office with public duties to perform merely on his passing a good examination is, I think, a bad mode of preventing mere patronage. My brother is one of the best generals that ever commanded an army, but the qualities that make him so are quite beyond the reach of any examination. Latterly the Cambridge examinations seem to turn upon very different matters from what prevailed in my time. I think a Cambridge education has for its object to make good members of society—not to extend science and make profound mathematicians. The tripos questions in the senate-house ought not to go beyond certain limits, and geometry ought to be cultivated and encouraged much more than it is."

Perhaps the Universities were spurred on by the influence of Edinburgh, the eighteenth century 'Athens of the North'. Perhaps they were inspired by Napoleon, who prided himself on his mathematics and expected it of his public servants, thus promoting the great French Écoles with Lagrange, Laplace, and all the others. Knowledge of their continental mathematics was brought to Cambridge around 1812 by Herschel, Baggage, and Peacock, influenced by Woodhouse, and it caught on. Their object was to 'do their best to leave the world wiser than they found it', and they worked with great determination. In 1817, Peacock wrote:

" · · · I shall never cease to exert myself to the utmost in the cause of reform, and I will never decline any office which may increase my power to effect it. · · · It is by silent perseverance only that we can hope to reduce the many-headed monster of prejudice, and make the University answer her character as

the loving mother of good learning and science." ([**3**] p. 121.)

Perhaps, too, the rise of nonconformism and the Evangelical Christians produced a reaction against sloth, but at any rate the competitive system, with pass-lists in order of merit, was soon in full swing at both the Universities, at least in the subjects of Mathematics and Greek. There were unfortunately few takers for the other subjects then offered. Some of those who were successful in this system entered the public service or became headmasters of public schools and began to spread this competitive system with enthusiasm.

A major effect on the public schools also was that the scholarship examinations were changed into their present form with a consequent rapid rise in the standards of work done. Unfortunately, the changes made it much harder for really poor students to go to Oxbridge, than in the eighteenth century, at least until the improvement in Secondary Education, which came much later. Also, Arnold of Rugby (1828–42) had made these schools morally acceptable. They had competition too from some of the new private schools established for the PB who had rising aspirations for their own sons, and rising incomes from trade to pay the fees. Elementary education was at first provided for the Working Class by voluntary organisations, but the first government grants to schools were made in the 1830's, and this created a need for an inspectorate and soon a need for the training of teachers. The first inspectors and later the first heads of training colleges were products of the competitive system at the two older Universities. On the whole they acted with a high-minded liberalism which has been of great benefit to the system, but which produced changes they did not foresee. As Montgomery argues in his book [**17**] they used the examination system as an administrative instrument for improvement in elementary schools, to allocate funds in a way which would avoid religious disputes with the organisations running elementary schools. Unfortunately, economics led to the disastrous system of 'payment by results,' which killed several promising experiments, presumably derived from Pestalozzi, Froebel, and the remarkable Hazelwood School established in Birmingham in 1819 (Spens [**19**] p. 20).

Meanwhile by mid-century the newly educated reformers had established the beginnings of a system of entry to the Civil Service by competitive examination as a remedy for the inefficiency caused by the system of patronage. (Purchase of Army commissions ceased in 1871.) They had to be tactful of course, because they were reflecting on the competence of existing office-holders, and also forging a weapon against the Aristocracy: *they* wanted the jobs! The system of examination spread and the schools responded by preparing to train boys for the examinations at different levels of the public and military service. For example, Manchester Grammar School had a 'Civil Service Form' in 1869; the school had been revived by R. F. Walker, one of the new products of the Honours degree system, who went in 1857, aged 27. It is presumably by these means that English Public Life has for so long been virtually free of corruption, a precious luxury not possessed by those other countries that inherited our political system in the 18th Century.

7. Social mobility. The 18th century Radicals and the French had talked of opening careers to the talented rather than the rich, on the ground of social justice. But to cut any ice, harder-headed grounds must be found, and the Prussians by 1870 had adopted the policy that efficiency was good for the State. Thus Matthew Arnold, an HMI, was sent to study the Prussian system, where an explicit object of the German *Gymnasium* was to provide highly educated servants of the State (copying from the Jesuits and the Chinese). Such a system tends to widen the reservoirs from which it draws its talent; it begins by accepting a static stratification of the society it is designed to run, but since it must expand it has an interest in talent rather than patronage. There is a consequent widening of its catchment area. Hence such a system carries unforeseen possibilities of social mobility, and then of a blurring of the class divisions. This method of constructing a 'Social Ladder' is a relatively cheap and quick way of producing a skilled Managerial Class in a backward nation. We shall consider its defects later on.

Another kind of mobility in society was that becoming available through the growth of the railways system. When white-collar jobs were advertised, candidates started to appear from different parts of England and it became necessary to compare the various qualifications that they presented. This was one of the reasons for the setting up of the Oxford and Cambridge Boards in the 1850's. Girls, too, were being thought of as possibilities for education(7) and after beginnings in the 1840's, leading to the founding in London of Queens and Bedford Colleges (1848, 1849), new High Schools for Girls were established after 1869 in large numbers. Newnham was founded in 1871 to prepare girls for the Cambridge Higher Local Examination, instituted in 1869, when Gerton was founded at Hitchin prior to moving to Cambridge in 1873. Everyone was rather surprised by the speed with which women teachers organised themselves and their institutions on the very latest (if possibly too academic) lines; they were far less prejudiced than men and far less bound by tradition (at least at that time). As an example of the emerging talent, Grace Chisholm was the first woman to receive a mathematics Ph.D. in Prussia (as related by Klein, [**13**] p. 179) in 1894, although much greater fame went to Phillipa Fawcett, who beat the Senior Wrangler at Cambridge in 1890, where she could not receive a degree. (Phillipa's family had much to do with getting Newnham started: see [**20**].)

The examinations were designed to test the kinds of mathematics thought to be appropriate to the three classes. For the Leisured Class, it was Euclid—for reasons relating to the Greeks and the notion of a gentleman(8) with some algebra and arithmetic. The difficulties naturally associated with the teaching of Euclid led conscientious schoolmasters to found our Association, with the initial aim of improving the Teaching of Geometry, later broadened to other branches of the subject. An unfortunate byproduct of its origin in the Leisured class, has been the relative neglect until recently, of other mathematical needs, such as led to the founding of the Association of Teachers of Mathematics (ATM) in the 1950's. Our modern timed examinations are very similar to those

designed for this system, themselves copied from the Oxbridge degree examinations.

It was the members of the Leisured class who planned the education of the Working class, and for vocational reasons offered mathematics in the form of arithmetic; and this against hostility to 'offering the 3R's to farm labourers,' reminiscent of today's cries of horror at having graduates for dustmen. The Forster Act caused a type of child to appear in the schools who needed even more basic material than arithmetic. Wardle [**23**] p. 43 quotes descriptions from contemporary writers, of children who had to be taught to sit without crouching, who expected a constant stream of blows, or who were like the inmates of Fagin's kitchen. Their teachers were regarded as artisans, not schoolmasters, having served apprenticeships as pupil-teachers, not as undergraduates. What a remarkable improvement do we see to-day, even with the defects we well know! But *we* are concerned with higher-order terms of the progression to perfection: *they* dealt with the zero-th order approximation, which, as so often with social problems, is the most difficult to get started yet produces astonishing dividends. As to examinations in these elementary schools, they were first carried out orally by inspectors, as in the Scottish Secondary Schools until much later; but presumably the growing numbers of children made it cheaper to adopt the timed exam (for specimens, see [**7**] p. 27 ff).

For the PB class, the schools varied from cheap private 'writing schools' on the one hand to the Grammar schools and some of the newer Public Schools on the other. It was natural for vocational reasons that this class should require Secondary Education for its sons. The need was reinforced by the increase in Elementary Education, for the pupil-teachers were felt by the Inspectorate to require further training, while rising expectations generated a class of pupils who simply wanted to stay at school and learn more. Great demands were also being made for more technical and scientific education, catalysed by the Great Exhibition of 1851, the founding of the Royal College of Science, the Royal School of Mines, and similar institutions. At the Exhibition, British products carried off most of the prizes, whereas at the Paris exhibition of 1867 foreign competition was superior; the Prussian victory in 1871 must have been like the Sputnik to the advocates of Science Education, as a spur for improving the quality of our manufacturing and engineering trade to compete with the Germans, Americans, and French. Thus, by the early 1900's there had grown a complicated system of Higher Grade Schools, Science Schools and Technical Schools all wanting to teach mathematics at a more advanced level, and from the point of view of Scientific applications. These schools were reorganised into the system of Grammar/Secondary schools following the 1902 Act, which raised common standards but perhaps eliminated some promising lines of growth. Of such schools, the Spens report [**19**] could note (p. 71) "their marked disinclination to deviate to any considerable extent from the main lines of the traditional grammar school curriculum. That conservative and imitative tendency which is so salient a characteristic in the evolution of English political and social institutions is particularly noticeable in this instance."

8. Curricula and Elitism. All these changes led to much discussion about the kind of mathematics the schools were offering. For the Leisured class, mathematics was chosen for a liberal education, but there was dispute as to whether it had the transference value claimed for it. Professor H. E. Armstrong, a chemist from the Royal College of Science and a leading exponent of 'free', pupil centred methods of teaching science, criticised the Leisure class concept of education as "suitable only for men who would spend £1000 a year and not for men who would earn £1000 a year." There was talk, too, of training a race of inventors; and argument about the value of the cultivated amateur compared with the narrow specialist. The differences in philosophy reveal themselves in arguments about examinations, centering on the difference between a competitive examination, designed to select the 'best' people (on some criterion) and a qualifying examination designed to throw out the 'incompetent' (on some other criterion). This linked itself to arguments about élitism and democracy, of the kind we are still hearing. There is the curious comment in the Gazette of 1909 that the effect of the new act was to bring in boys on free places and thus 'to drive away' the kind of pupil that it was thought desirable to have in those days—those free places being probably responsible for the appearance of most of the audience here today. On the other hand, Sir George Greenhill, a great élitist of a very old fashioned kind, suspicious of the French, gave his Presidential address to the Association in 1912; and we have a good insight into his beliefs with the following quotation (Gazette, VII p. 257) in spite of its contradictory final sentence:—

"Forty years ago it was felt that the Cambridge School of Mathematics was too provincial, and outside examiners were called in—Maxwell, and Thomson,[9] and Tait. A wonderful stimulus was the result on the enlargement of the outlook, in those the palmiest days.

"But the effort died out with the disastrous division of the Tripos into two parts. The democracy got all it wanted in part one, and the second part died of inanition. A dreadful system has been copied everywhere, from Oxford, of three classes in alphabetical order; so that two-thirds of the candidates for Honours, however few, are dishonoured for life with a second and third class.

"If order of merit is to go, abolish the class as well; so that a man who fails does not appear in the list. It is easy to award a mark of distinction in a subject.

"But the examiner, prosaic and conscientious, considers himself tied down by his instructions, and is determined to have a third class in his list, however small, although the third class man may think he has stopped away.

"This is called raising the standard, up to infinity. At the same time democracy is engaged in its occupation of lowering the standard for the mass, by cutting off one year of study, and that the most valuable. All achieving the same result in a different manner."

Those who could discuss the changes that might be made in mathematics courses for the average chap were most likely to be either people from the Universities who knew about Euclid and about Calculus, or technically-trained

people who wanted mathematics for applications, or those who taught teachers in training colleges to prepare them for the teaching of what we know as Secondary Mathematics. These people must have been greatly preoccupied by the changes from 1870, and they do not seem to have written much until the 1890's, when they began to review their labours, for example in the Gazette. From their side-remarks, one can infer things about their own youth, but they are largely representative of the Leisure class.

9. Some spokesmen. A wonderful example of the first class is Canon J. J. Wilson who gave a presidential address to the Mathematical Association exactly fifty years ago. He must have been quite old by then[(10)] and his reminiscences are very interesting; one or two other people added their own reminiscences at the end of his printed lecture in the Gazette (Vol. X, p. 239). When they were boys (say in the 1860's) they were seeing the last of an old kind of schoolmaster who was ignorant and lazy, and meeting the new kind of bustling keen Senior Wrangler, such as Wilson himself became before he returned as a mathematics master to Rugby. By 1867 at the suggestion of Temple, he had written the first English text-book which departed from the style of Euclid, thereby acquiring some fame. He recalls how he was invited to lecture in Edinburgh in 1870 at the Royal High School, where he was presented with an illuminated address, signed by many distinguished citizens of Edinburgh and other Scotch cities, including the Lord Provost, MP's, Professors, Headmasters, and Literati. They were thanking him for his efforts to change the teaching of Geometry in which they all had such a vital interest; we must remember this was Scotland where interest in education was rather different from what it was in England. The average man there was quite different too!

Canon Wilson then goes on to tell an amusing story about a friendship he struck up with boys from Manchester Grammar School when travelling on the train between Manchester and Rochdale at the time he was vicar there. They boasted to him of a marvellous mathematician in the school and one could imagine a wonderful T.V. sketch of the sort of thing that got said in a juxtaposition of accents. He sent some problems back for this bright mathematician to do, and in desperation at the boy's success Wilson took an unfair advantage and eventually set him the problem of showing that every prime of the form $4n+1$ is a sum of two squares (a fact now treated experimentally in contemporary primary schools!). He knew it was hard, but it is interesting that he seemed to think of it as a problem on the same footing as those he had offered previously (on the sums of binomial coefficients, etc.). The bright boy admitted defeat, and his schoolmasters could not help him either, so Wilson felt that he had kept his dignity as a Senior Wrangler. His presidential address allowed him to meet that bright school boy of 30 years previously and it turned out to be Sir E. T. Whittaker, the retiring president of the Mathematical Association.

As an example of the second kind of reformer, with a scientific-technical interest (but a leisure class background), we might look at Professor J. Perry,

who in the 1880's taught Science at Clifton (one of the new Private schools). His reforming ideas were supported by this Association. He reveals much of himself in his address of 1909 (Gazette V, p. 1) to a conference held in conjunction with the Federated Associations of London Non-Primary Teachers (what hierarchical nuances that phrase reveals!). The subject was 'The correlation of the teaching of Mathematics and Science', and Perry gives a lecture followed by a discussion. Though energetic, he was not the subtlest of men, but a great one for slogans—'A doubling of salaries, halving of classes, the ousting of mere specialists, and completely getting rid of the outside examiner.' His general practical approach shows him as a precursor of the Nuffield Project, and his own primary education sounds (p. 11) wonderful, though it would offend our Black Paper Writers. The knowledge of mathematics that he possesses is very much that of the educated average chap that we could well have in mind when we aim to have universal numeracy—based on calculus and common sense. In the discussion following Perry's lecture, there are many enlightened views whose general tenor is to add a mathematician's gloss and tact to Perry's proposals. One voice, that of Dr. Percy Nunn, represents the Training Colleges, with an interesting teaching suggestion, although he claims to be no mathematician. For an authentic 'Working Class' view, however, it would be desirable to know what was happening in the Mechanics' Institutes of the century. Certainly working men existed with deep learning, but one hears of them only via anecdotes told by members of the Leisure Class who came across them; for a Mathematical Carpenter see Gazette, Vol. XXXII. p. 133.

Nunn's views appear again in a set of reports [**7**] published in 1912 by the Board of Education, following a request from the 1908 International Congress of Mathematicians, for gathering a survey of the state of mathematical education in different countries. These reports are extremely interesting, and many of them are by well-known text-book writers and members of this Association, like Godfrey, Siddons, Barnard, Durrell, Gibson, Hardy, etc. There are papers, too, on Arithmetic in Elementary Schools.

There is evidence here of an enormous amount of activity and excitement because so many of the people were involved in planning new courses for the new demands of the time. Recall that in 1902 the Secondary Education Act had been passed, to which so many modern Grammar Schools owe their continued existence. There was increasing pressure for technical education. The Navy, strangely enough, had only recently accepted competitive examinations for entry, and this seems to have given a great impulse to the Naval Schools at Osborne and Dartmouth, judging by a model account of the planning of a mathematical curriculum with all the appropriate constraints, given by Mercer of Dartmouth.

Reading these reports one wonders what happened to Geometrical and Mechanical Drawing. There are several enthusiasts for it in these papers, and a very exciting course is outlined by D. M. Lowe. Nowadays one can get high marks in A-Level Mathematics without doing anything outside the plane, and the only three-dimensional thinking now encouraged is that for the (single) solid

trigonometry question. Our contemporary papers in engineering drawing involve far more real mathematical imagination than the present traditional A-Level papers, yet they are much disdained by Schools and Universities. Why were Lowe and his fellow enthusiasts ignored?

These articles indicate, clearly, that the average active member of the profession around this period was then very much like the average active member now. He cared deeply about his job and he probably knew roughly the same amount of mathematics; although, for example, Barnard writes a very amusing piece about rigour in mathematics, obviously having been imbued with it by Hardy and the younger men at Cambridge, whereas Siddons confesses that he never learned rigour and is against the 'mental-discipline' school of mathematics teaching.

10. The blight of examinations. But, in spite of the excitement, a fairly common theme runs through these papers, as with Tait,(11) Armstrong, Perry, and many others. It concerns the blighting effect of the examination system, at least with externally imposed examinations. Teachers, they said, were already teaching only for the examinations, using blind rote-learning; and good teachers were prevented from doing many of the things that they would wish because their pupils might fail examinations.(12) True, they are pleased that several of their recommendations about improvements in the syllabus were being accepted within the system; for example, it is now no longer necessary to teach Physics without Calculus, although around 1900 few mathematicians knew any science. The curious air of impotence against the examination system possibly derives from an over-developed high-minded liberal attitude about fairness, as exemplified by one of the reminiscences (Gazette, Vol. X, p. 246) following Canon Wilson's, where a resolution, stating what examiners should do, had to be modified on the grounds that 'it would mean dictating to examiners.' The effect of that impotence can be judged from the examination papers that are included with these 1912 reports. They are depressingly similar to our present traditional papers (although one HSC question set by Bristol in 1919 was: 'In what sense is $\sqrt{2}$ a number?'—quite banned nowadays). Curiously, other subjects *have* changed the quality of their examinations within living memory. Why can't mathematics?

As I see it, our problem as professional teachers of mathematics is to be able sensibly to link the growth of the mathematical organism with the changing needs of our Society. Thus we need to be able (a) to find out, as individuals, how mathematics itself is changing and why; (b) to select and digest material appropriate to our purposes as teachers (and reject it if necessary); then (c) to work out ways of *communicating* it.

It is not a simple matter of 'modernising' courses in the sense of adding new material; many of our Universities have done just that, and yet one has only to ask for independent thought from Honours students with high marks, to echo the sentiments of that Elementary School Headmaster of 1912 who said 'It does not necessarily follow, because children reproduce the language of reasoning,

that they have reasoned'. Much more is it a matter of adapting our *style* of teaching to be appropriate both to the kind of mathematics and the kind of pupil we happen to be working with. Similarly, different situations need different ways of assessment. But none of these changes in techniques can be worked out unless we have freedom to make teaching mistakes (in order to learn from them). Such freedom is impossible within a system of formal examinations, if they are externally imposed. Now, society is clearly not static: our education system alone has had constantly to adapt to change over the past century or more, but within a framework *designed* for a static Mandarin society. If we want to be real professionals, we have to challenge that static concept, just as the Medical Profession and others have told society what is good for it. We have to remove the often-stated belief that all the things we do in Education have long traditions behind them: few have, but one of these is the tradition of examining in the conventional way. That system was developed, first to relieve the boredom of rich but clever undergraduates (to give them a competition) and then to defeat the system of patronage, as a counter to human greed. Neither of these purposes has any necessary connexion with education in spite of well-known philosophical arguments about goals and competition (see [**2**] p. 121). The high-minded sense of fairness implicit in the present examining system must be retained, for sure, but it is time we developed a system that was fair to our *mathematics* as well as to our students.

For example, it has taken almost a hundred years for such notions as matrices and congruences in number theory to come into common mathematics. It has taken this time, also, for many of the technical parts of Riemann's theory of functions and the related theory of Fuchsian groups as developed by Klein and Poincaré to become well understood with properly constructed proofs. With this understanding has come some reorganisation of mathematics and the realisation that by using the axiomatic method one can unify many parts of mathematics, without spending a long time on details. This 'Renewal principle' of mathematics has always been present from the time of the Greeks and presumably before; but it has never previously hit mathematicians so hard and so quickly, because of the recent rapid increase in mathematical activity throughout the world. Naturally, it has to be reflected in teaching at all levels for reasons which are both practical and aesthetic.

We need to concentrate on the dynamic aspects of mathematics rather than the static ones because the dynamic ones are related to contemporary creativity (both in mathematics and in other spheres of thought). Also if we teach mathematics because the bulk of our pupils will need to *use* it, then it is no good teaching frozen examination pieces, without relation to anything else except passing examinations. They need training for a life and jobs that may well change out of recognition within the span of their working lives. Since we want children to understand rather than remember, it is natural to use methods of the kind advocated by the best kindergarten teachers over the past century. This I have called [**10**] the 'open tendency' in the teaching of mathematics. It is pupil-

centred and our Black Paperers do not understand it, but it seems to be a better reflection of the way that real mathematics is created by argument and discussion, shunning dogmatic, authoritarian assertion. This search for the truth by inquiry, so beautifully explained in the works of Bertrand Russell, has kept our subject young in spite of its great age. If we are to keep its appeal to the young, then this method, which implies a fraternity of our craft, must be kept to the fore.

All these considerations have been in the minds of our contemporary reformers, and grossly misunderstood by those for whom mathematics is a static agility. These considerations were already clearly set out, and argued with profound insight, in the Presidential Address of A. N. Whitehead *fifty-five years ago*, (and reprinted in Gazette, Vol. XXXII, (1948) p. 110). Although his argument is largely aesthetic and moral (and hence likely to be ignored), he includes a paragraph (p. 119) of warning from the point of view of the British Nation's survival, though in terms too general to be understood at the time, and in the middle of the Great War. However, his introductory sentence to the paragraph is worth quoting, if only to show Whitehead's humanity:—"When one considers in its length and in its breadth the importance of this question of the education of a nation's young, the broken lives, the defeated hopes, the national failures, which result from the frivolous inertia with which it is treated, it is difficult to restrain within oneself a savage rage."

Nowadays we are no longer so strongly bound by those social inhibitions, inherent in the rise up the 'Social Ladder', that may have stopped Tait and the others from challenging the system by deeds rather than words. For we have in England three techniques already available, though requiring improvement and development (and Whitehead's Address suggests lines of attack):

(i) the CSE modes of examination, involving partial course-work (in 'Mode 1') or timed examinations set and marked by the teacher and associates ('Mode 3');
(ii) the possibility of Project work in Mathematics A-level now allowed by the Southern Universities Joint Board;
(iii) the ability to band together and work on such schemes as SMP, MME, MEI, etc. (These schemes have not, of course, escaped the bad effects of the conventional mode of examination.)

We cannot abolish the conventional system overnight, and economic considerations will, in any case, force us to use our assessments in order to select pupils for scarce places. We shall have to face the problem of patronage (or favoritism) that our present examination system has circumvented. I do not see this problem as an insuperable obstacle, and in any case, we *must* overcome it if we want a substitute for the old system. I do not deny that this will take a lot of hard work and co-operation from us all. We shall meet particularly strong opposition from those in our society who have done well out of conventional examinations to climb the 'Social Ladder', and who consequently think that

Education is largely something we change our status with. Several of our politicians climbed the ladder that way, and their speeches show that it is the climb for them that counts, not Truth or Creativity or fulfillment of personality. Many voters agree with them, of course.

However, if we do not overcome those atavistic feelings, if we continue to accept the blighting effect of examinations, then we cannot be a real, self-regulating professional organisation. And if we do not, then I believe, pessimistically, that the average chap of 2071 will be little better off mathematically than his counterpart today.

NOTES

(1) In the investigation I was greatly helped and guided by colleagues in the Mathematics and Education Departments of the University of Southampton.

(2) Or rather, from the Gaussian sphere of complex numbers together with the 'point at infinity'.

(3) Discussed in the book of Moise [**16**] p. 383.

(4) The 'Black Papers' form collections of essays, all critical of 'modern' trends in education, including the recent adoption of schemes for Comprehensive schools in Britain. With some exceptions the standard of criticism is poor, but the Papers represent a strong political backlash that cannot be ignored.

(4a) References to 'Gazette' are to the official Journal of the Mathematical Association (the 'Mathematical Gazette').

(5) Robertson Smith, see [**14**] p. 291. His Aberdeen professorship was in the Free Church College. With Tait, he *initiated* the teaching of experimental physics in Edinburgh between 1868 and 1869.

(6) As a sign of the increasing academic respectability of Medicine, the Manchester Royal School of Medicine was incorporated into Owens College (University of Manchester), in 1872.

(7) For more details, see the Spens report [**19**] p. 42 ff.

(8) On Veblen's theory [**22**], facility in Euclid would be a good way of displaying 'conspicuous consumption', and allaying suspicion that one might have had to spend time in conventional work. But, for an interesting semi-counterexample, see the account [**8**] of Dawson, ex-shepherd boy trainer of Wranglers in the 18th century.

(9) Thomson later became Lord Kelvin.

(10) I am grateful to Mr. J. T. Combridge for the information that Wilson was 85 at the time, and lived a further ten years. He was Canon of Worcester Cathedral.

(11) See 'The evils of cram' [**14**] p. 249 where Tait's conservative political attitudes are contrasted. For a devastating attack on the Tripos examination, see the Presidential Address of 1926 by G. H. Hardy 'The case against the Tripos', reprinted in Gazette, Vol. XXXII p. 134, having had no effect in the intervening 22 years (or the following 23).

(12) Armstrong's pupils did fail; see Richmond and Quereshi [**18**] p. 514.

References

1. R. Abraham, Foundations of Mechanics, Benjamin, New York, 1967.
2. W. H. G. Armytage, Four Hundred Years of English Education, Cambridge, 1965.
3. W. W. R. Ball, A History of the study of Mathematics in Cambridge, Cambridge, 1889.
4. E. T. Bell, Men of Mathematics, Penguin Books (1953).
5. ———, The Development of Mathematics, McGraw-Hill, New York, 1945.
6. Black Papers, Critical Quarterly Society Publications, London, 1969–70.
7. Board of Education Report, The Teaching of Mathematics in the U.K., London, 1912.
8. C. C. Booth, John Dawson (1734–1820), British Medical J., Oct. 1970.

9. A. H. Clifford, The Common Sense of the exact sciences, Dover Reprint, New York, 1955.

10. H. B. Griffiths, Mathematical insight and mathematical curricula, Proc. Int. Cong. Math., (1970) (to appear).

11. H. B. Griffiths and P. J. Hilton, Classical Mathematics, Van Nostrand, Princeton, N. J., 1970.

12. J. Hammersley, On the enfeeblement of mathematical skills by the teaching of modern mathematics and similar intellectual trash in schools and Universities, Bull. Inst. Math. and Applics., 4 (1968) 68–85.

13. F. Klein, Elementary Mathematics from the advanced standpoint, Dover, New York, 1939.

14. C. G. Knott, The Life and Scientific Work of P. G. Tait, Cambridge University Press, 1910.

15. K. O. May, Quantitative Growth of the Mathematical Literature, Science, 154 (1966) 1672–1673.

16. E. E. Moise, Elementary Geometry from an advanced standpoint, Addison-Wesley, Reading, Mass., 1963.

17. R. J. Montgomery, Examinations. An account of their evolution as administrative devices in England, Longmans, London 1965.

18. A. R. Quereshi and P. E. Richmond, Armstrong's heuristic method in 1964, School Science Review, No. 157, June 1964.

19. Spens Report, Report of the Consultative Committee on Secondary Education, with special reference to Grammar Schools and Technical High Schools, Board of Education, London 1938.

20. R. Strachey, Millicent Garrett Fawcett, Murray, London ,1931.

21. D'Arcy W. Thompson, On Growth and Form, Abridged and edited by J. T. Bonner, Cambridge, 1962.

22. Thorstein B. Veblen, The Theory of the Leisure Class, Allen and Unwin, London, 1918.

23. D. Wardle, English Popular Education 1780–1970, Cambridge 1970.

SIGNS OF DERIVATIVES AND ANALYTIC BEHAVIOR

R. P. BOAS, Jr., Northwestern University

This is an account of some striking results, most of which are far from new but not widely known. Although many of them were quite unexpected when they were discovered, the results themselves are easily comprehended by undergraduates; also, many of the proofs are sufficiently elementary to be presented in a course in advanced calculus or elementary real analysis, or in an under-

---

Ralph Boas did his Harvard Ph.D. under D. V. Widder; he was an NRC Fellow at Princeton and Cambridge Universities and Instructor at Duke and the USN Pre-Flight School, Chapel Hill and Lecturer at Harvard and M.I.T. He served as Editor of Mathematical Reviews for five years before joining Northwestern, where he is presently H. S. Noyes Professor and Mathematics Department Chairman. He has held Guggenheim and Northwestern President's Fellowships, has served as Vice-president and Trustee of the AMS, Vice-president and Chairman of Section A, AAAS, Chairman of CUPM, and has chaired many AMS and MAA committees. His main research is in function theory, and he is the author of *Entire Functions* (1945), *Polynomial Expansions of Analytic Functions* (with R. C. Buck, 1958), *Primer of Real Functions* (Carus Monograph, 1960), and *Integrability Theorems for Trigonometric Transforms* (1967). *Editor.*

graduate seminar. The article is intended as a "resource paper," rather than a formal exposition, and accordingly I have omitted proofs that can be found in easily accessible sources.

1. Derivatives all positive. We are concerned with real functions that have derivatives of all orders. The field we are considering began in 1914, when S. Bernstein proved that if $f^{(k)}(x) \geqq 0$ for all x on $[a, b]$, then f is real-analytic, in fact is the restriction of a function that is analytic in a disk centered at a and of radius $b-a$. (We shall usually disregard the distinction between an analytic function and its restriction to the real axis, and simply say "f is analytic in a disk" in this case.) The rather simple proof is reproduced in a number of places, for example [24], p. 146; [6], p. 155. (It is somewhat harder to show that it is enough just to assume that $f^{(k)}(x) \geqq 0$ for $k \geqq n(x)$, where $n(x)$ may depend on x.)

A function with all derivatives nonnegative is called **absolutely monotonic**. A function whose successive derivatives alternate in sign, so that $(-1)^n f^{(n)}(x) \geqq 0$, is called **completely monotonic**; the change of variable $y=b+a-x$ converts a member of one class into a member of the other. Naturally a completely monotonic function is analytic in a disk centered at b.

Many of the familiar functions that occur in calculus are either absolutely or completely monotonic, and Bernstein's theorem then provides an immediate proof that they are represented by their power series. Obvious examples are e^x, e^{-x}, and $(1-x)^{-1}$. Although $(1-x)^r$ is not necessarily absolutely monotonic, one of its derivatives of sufficiently high order is so, and we obtain an easy proof of the binomial theorem for general real exponents. On the other hand, although $\tan x$ is absolutely monotonic on $[0, \pi/2)$, it would not be easy to establish this by direct inspection of the derivatives of $\tan x$.

A function that is absolutely monotonic on $[0, \infty)$ is the restriction of an **entire function** (one that is analytic in the whole finite complex plane). On the other hand, when a function is completely monotonic on $[0, \infty)$, as are $1/(x+1)$ and e^{-cx} $(c>0)$, the statement about where the function is analytic has to be modified; what is in fact true is that the function is analytic in a right-hand half-plane. A little thought shows that a function defined by a convergent Laplace integral of the form

$$f(x) = \int_0^\infty e^{-xt}\phi(t)dt, \qquad \phi(t) \geqq 0,$$

or more generally by a convergent Laplace-Stieltjes integral

$$f(x) = \int_0^\infty e^{-xt}d\alpha(t), \qquad \alpha \text{ increasing,} \tag{1}$$

is completely monotonic where it converges. The converse is also true: Bernstein and Widder discovered independently about 1929 that *every* function completely monotonic on a half-line (a, ∞) is a Laplace transform of the form (1). There

is a full account of the subject (which is neither elementary nor wholly germane to this article) in [**24**].

It is interesting, and sometimes useful, to know that a function, initially known only to be continuous, is absolutely monotonic if all its differences are nonnegative, that is,

$$\Delta_h^k f(x) = \sum_{j=0}^{k} (-1)^j \binom{k}{j} f(x+jh) \geqq 0 \tag{2}$$

for all nonnegative integers k, for all x and all positive h such that the points $x+jh$ that occur in (2) are in the domain of f ([**24**], Chap. 4). Actually it can be shown that (2) for $k \leqq n$ makes f have continuous derivatives of orders up to and including $n-2$; this is quite elementary, but not entirely trivial [**9**].

Bernstein's theorem on absolutely monotonic functions has been extended to functions with domain of dimension greater than 1 [**19**], and even to functions with infinite-dimensional domain [**22**].

2. Each derivative has a fixed sign. Perhaps the most natural next step is to consider functions for which each derivative is of fixed sign on $[a, b]$, without regard to how the signs are distributed. Bernstein did this; he called such functions **regularly monotonic,** and showed that a regularly monotonic function is always analytic ([**1**], pp. 196–197). However, the function does not have to be analytic in as large a region as in the absolutely monotonic case. Bernstein's proof is a remarkable application of the elements of the theory of the approximation of continuous functions by polynomials; since there does not appear to be any readily accessible account of it, I reproduce it here, giving the necessary background in an appendix.

LEMMA 1 (see p. 1090). *If $f^{(n)}(x) \geqq N > 0$ on an interval I of length $2h$, and M is the maximum of $|f(x)|$ on I, then $M \geqq 2\,N(h/2)^n/n!$.*

Suppose now that $f^{(n+1)}(x)$ has constant sign on I; then $f^{(n)}(x)$ is monotonic (either increasing or decreasing). Let t be any point of I; we may suppose $f^{(n)}(t) > 0$ (otherwise consider $-f(x)$). Since $f^{(n)}$ is monotonic there is either an interval $(t, t+\epsilon)$ or an interval $(t-\epsilon, t)$ on which $f^{(n)}(x) > f^{(n)}(t) > 0$ (where ϵ can be taken as the distance from t to the nearer endpoint of I). Hence by Lemma 1

$$M > 2\,|f^{(n)}(t)|\,(\epsilon/4)^n/n!$$

for each t in I. That is,

$$|f^{(n)}(t)| \leqq \tfrac{1}{2} n!\, M(4/\epsilon)^n. \tag{3}$$

If we expand $f(x)$ in a Taylor series about a point s of I, the usual estimate for the remainder after n terms yields

$$|R_n| \leqq |x-s|^n \max |f^{(n)}(t)|/n!,$$

where the maximum is taken for t between s and x. If then $|x-s|$ is less than one-fourth the smaller of the distances from s and from x to the endpoints of I, we obtain $R_n \to 0$.

Bernstein next asked what happens if not all $f^{(n)}(x)$ have constant sign on I, but infinitely many of them do. He showed that f then always has a "quasi-analytic" property, namely that f is determined throughout I by its values on an arbitrarily short subinterval. If enough of the $f^{(n)}$ have constant sign, f is still analytic; more precisely, this happens when $f^{(n_k)}(x) \geqq 0$ with n_{k+1}/n_k bounded; for example, if $f^{(2k)}(x) \geqq 0$ or even if $f^{(2^k)}(x) \geqq 0$, but not if $f^{(k!)}(x) \geqq 0$. (See [**4**].)

3. Sequence of derivatives of fixed sign. These results of Bernstein's are far from simple to establish. In the summer of 1940, Widder asked me if I knew a simple proof that f is analytic when $f^{(2n)}(x) \geqq 0$. I was not immediately able to provide one, but I suggested that he try to use Lidstone series, which are series of polynomials, two of each odd degree, with coefficients $f^{(2n)}(1)$ and $f^{(2n)}(0)$. This seems, in retrospect, not a very sensible suggestion because what is "really" involved in Lidstone series is $(-1)^n f^{(2n)}(1)$ and $(-1)^n f^{(2n)}(0)$. It did not, in fact, lead to a proof of the theorem in question; but it led to the quite unexpected result that if the even derivatives of f alternate in sign on (0, 1), i.e., if $(-1)^n f^{(2n)}(x) \geqq 0$, then f is represented by a Lidstone series and consequently is not only the restriction of an analytic function, but of one that is entire and of slow growth. (See [**24**], pp. 177–179.) More precisely, it satisfies $|f^{(n)}(x)| \leqq A\pi^n$, where A does not depend on n, and hence is what is known as an entire function of **exponential type,** satisfying $|f(z)| \leqq Ae^{\pi|z|}$. A function f with $(-1)^n f^{(2n)}(x) \geqq 0$ in an interval is now called **completely convex.** The discovery of completely convex functions led immediately to a few years of intense development of related results, after which the field became rather inactive, although a number of open problems remain.

There is an elementary proof of Widder's theorem ([**24**], p. 177). It involves repeated integration by parts in

$$\int_0^1 f(x) \sin \pi x dx;$$

this leads to the necessary estimates, but only for even n. For the transition to odd n, one needs a lemma of Hadamard's:

If on $[-h, h]$ we have $|f(x)| \leqq A$ and $|f''(x)| \leqq B$, and $B/A > 4/h^2$, then $|f'(x)| \leqq 2(AB)^{\frac{1}{2}}$.

Proof: By Taylor's theorem with remainder of order 2,

$$f'(x) = \frac{f(x+\delta) - f(x)}{\delta} - \frac{1}{2}\delta f''(x+\theta\delta), |\theta| < 1; \; |f'(x)| \leqq 2A/|\delta| + |\delta| B/2.$$

If $B/A > 4/h^2$ we can take $\delta = 2(A/B)^{\frac{1}{2}}$.

Incidentally this lemma is one of a family of results in which one infers something about f' from information about f and f''; some references are [**14**], p. 36; [**3**]. Many generalizations of completely convex functions rely on more general (and much deeper) inequalities of the following form: Let $M_k = \max |f^{(k)}(x)|$; then if $M_n \neq 0$, there are numbers $C_{n,k}$ such that

$$|f^{(k)}(x)| \leqq C_{n,k} M_0^{1-1/n} M_n^{k/n} \qquad (0 < k < n);$$

fairly precise estimates for $C_{n,k}$ are required. At present, the $C_{n,k}$ are known explicitly for functions whose domain is $(-\infty, \infty)$ [**12**], and for those whose domain is $[0, \infty)$ [**20**]; various estimates are known when the domain is a finite interval ([**8**], and references given there).

The idea of the elementary proof of Widder's theorem was used by Pólya [**15**] to show that f is an entire function of exponential type if $f^{(k)}(x) \sin (k+1)\alpha \geqq 0$ for some α $(0<\alpha<\pi)$ and all k; it seems that more general results of this kind have not been studied. The elementary proof, however, does not work even for such a regular case as $(-1)^k f^{(4k)}(x) \geqq 0$, and it does not give any insight into why completely monotonic and completely convex functions behave so differently.

At this point we should observe something that had been forgotten in 1940, namely that about 1928 Bernstein had already found that the distribution of the signs of successive derivatives has a decisive influence on the behavior of a regularly monotonic function. More precisely, the significant property in his work is the distribution of successive blocks of either constant or alternating sign. If each derivative has a fixed sign, each derivative is monotonic; the property is easier to state in terms of whether $|f^{(n)}(x)|$ increases or decreases (as was suggested by Pólya). Since the derivative of $[f^{(n)}(x)]^2$ is $2f^{(n)}(x)f^{(n+1)}(x)$, we have $|f^{(n)}(x)|$ increasing if $f^{(n)}(x)$ and $f^{(n+1)}(x)$ have the same sign, $|f^{(n)}(x)|$ decreasing if $f^{(n)}(x)$ and $f^{(n+1)}(x)$ have opposite signs. We consider successive blocks where $|f^{(n)}(x)|$ increases or decreases; for example, when the signs of successive derivatives are $+++-+-+----+-+-+++$, etc., the lengths of the successive blocks are 2, 5, 2, 5, etc. For $\sin x$ on $(0, \pi/2)$, the signs are $++--++--$, etc., and all blocks are of length 1. For an absolutely or completely monotonic function, there is just one block, of infinite length. A convenient way to see where one block ends and the next begins is to notice that $f^{(n)}$ and $f^{(n+1)}$ belong to different blocks if and only if $f^{(n)}(x)f^{(n+2)}(x) < 0$. (See [**16**], p. 185.) (Functions with the signs of successive derivatives distributed periodically, for example like those of $\sin x$ on $(0, \pi/2)$, are called **cyclically monotonic**; they have a substantial literature of their own.) The general lesson of Bernstein's results (which will be stated in greater detail below) is that the presence of many blocks makes the function behave regularly, and more regularly when the blocks are short.

However, Bernstein's results do not establish Widder's theorem since that theorem has no hypothesis at all about the derivatives of odd order. The follow-

ing general theorem, which includes both Bernstein's and Widder's results, appeared two years after Widder's theorem [**8**].

Let $\{n_k\}$ *and* $\{q_k\}$ *be increasing sequences of positive integers such that* $q_1+q_2+\cdots+q_k=O(n_k)$, *and suppose that* $f^{(n_k)}(x)$ *and* $f^{(n_k+2q_k)}(x)$ *have opposite signs (so that it is derivatives with orders differing by an even integer that have opposite signs). Then if:*

$n_k-n_{k-1}=o(n_k)$ *and* $q_k=o(n_k)$, f *is entire;*
$n_k-n_{k-1}=O(n_k^{(\rho-1)/\rho})$ *and* $q_k=o(n_k)$, f *is entire and of order at most* ρ;
$n_k-n_{k-1}=O(1)$ *and* $q_k=O(1)$, f *is entire and of exponential type.*

For example, when $n_k=2k$ and $q_k=1$ we have completely convex functions; when $n_k=4k$ and $q_k=2$, we have functions such that $(-1)^kf^{(4k)}(x)\geqq 0$; when $l_1+l_2+\cdots+l_k=n_k$ (l_k is the length of the kth block for a regularly monotonic function), and $q_k=1$, we get Bernstein's results on blocks of signs. When $n_k=k^2$ and $q_k=k$ the theorem says that if $f^{(k^2)}(x)$ and $f^{((k+1)^2-1)}(x)$ always have opposite signs, for example if $f'(x), f^{(4)}(x), f^{(9)}(x), \cdots \geqq 0$ and $f(x), f'''(x), f^{(8)}(x), \cdots \leqq 0$, then f is an entire function of order at most 2.

It is interesting that in spite of the apparent generality of this theorem, there still are theorems with simple statements that it does not cover. Indeed, Leeming and Sharma [**13**] have shown that f is entire and of exponential type if $(-1)^kf^{(pk)}(x)\geqq 0$ and $(-1)^kf^{(pk+l)}(a)\geqq 0$ for $l=1, 2, \cdots, p-2$. Note that nothing is said about the sign of $f^{(pk-1)}(x)$, and nothing about the intermediate derivatives except at one point; we are again outside the domain of Bernstein's results. Leeming and Sharma base their proof on generalized Lidstone series; it would be interesting to have a direct elementary proof.

So far we have dealt with functions such that each derivative, or each of a sequence of derivatives, has no zeros on an interval. Suppose instead that no derivative changes sign more than a prescribed number of times, say that $f^{(n)}(x)$ has at most N_n zeros. It is reasonable to suppose that f will be more well-behaved when N_n is small. In fact, in 1943 Schaeffer [**18**] showed that if N_n is bounded for $a\leqq x\leqq b$, then f is analytic there. A considerable number of further results were obtained between 1940 and 1943; see [**16**]. Since 1943 the field has been rather inactive. See, however, [**17**] for some generalizations of completely convex functions, and [**5**] for a representation of completely convex functions. There are still a number of open questions. For example, suppose that $f^{(n_k)}(x)$ has at most N zeros for a sequence $\{n_k\}\neq\{k\}$? Suppose that $f^{(k)}(x)$ has at most $N(k)$ zeros? Suppose that $f^{(n)}(x)\geqq 0$ only in an interval I_n, where the length of I_n does not decrease too fast? Completely monotonic sequences $\{\mu_n\}$ are defined by having their differences of alternating sign; they have a complete theory (see [**24**], Chap. 3). Nothing seems to be known about completely convex sequences.

Appendix. *Proof of Lemma 1.* Except for Chebyshev's theorem on best approximation, the proof is entirely elementary, although rather exacting. Our version is expanded from the outline given in [**1**], pp. 8–10.

LEMMA 2 (Chebyshev's theorem). *If f is a real continuous function on a finite interval $[a, b]$, there is a unique polynomial P_n of degree at most n that approximates f most closely on $[a, b]$, in the sense that $\max_x|f(x)-P_n(x)|$ is as small as possible. The minimum is denoted by $E_n[f]$. The polynomial P_n is characterized by the property that $f(x)-P_n(x)$ has at least $n+2$ extrema on $[a, b]$, where $|f(x)-P_n(x)| =E_n[f]$ and the signs of $f(x)-P_n(x)$ alternate at successive extrema.*

A proof of Lemma 2 can be found in almost any book on the theory of approximation, for example [**10**] or [**21**]. The proof of Lemma 1 exploits all the information furnished by Lemma 2.

LEMMA 3. *If $g^{(n+1)}(x)$ is continuous and strictly positive in $[a, b]$, and $g(x)$ has exactly $n+1$ changes of sign in (a, b), then $g(b)>0$.*

Since $g(x)$ has $n+1$ changes of sign, $g'(x)$ has at least n, $g''(x)$ has at least $n-1$, and so on. Finally, $g^{(n)}(x)$ has at least one, and it cannot have more, since $g^{(n+1)}(x)$ has none. Let $g^{(n)}(x)$ change sign at y_1, where of course $g^{(n)}(y_1)=0$. If $x>y_1$,

$$g^{(n)}(x) = \int_{y_1}^{x} g^{(n+1)}(t)dt > 0.$$

Similarly, $g^{(n-1)}(x)$ has two changes of sign, say at z_1, z_2 with $z_1<y_1<z_2$; and so, for $x>z_2$, $g^{(n-1)}(x)>0$. This clearly starts an induction that winds up with $g(b)>0$.

LEMMA 4. *If ϕ and f have continuous $(n+1)$-th derivatives on $[a, b]$ and $0<\phi^{(n+1)}(x)<f^{(n+1)}(x)$ on $[a, b]$, then $E_n[\phi]<E_n[f]$.*

Let P_n, Q_n be the polynomials of degree n of best approximation to ϕ and f, respectively; write $D_n(x)=\phi(x)-P_n(x)$. Then D_n has at least $n+2$ extrema with alternating signs and so changes sign at $n+1$ points at least. If D_n had more than $n+1$ changes of sign, $D_n^{(n+1)}$ would have at least one, but $D_n^{(n+1)}(x)=\phi^{(n+1)}(x)>0$. Hence D_n has exactly $n+1$ changes of sign and D_n' has exactly n; and D_n has exactly $n+2$ extrema on the closed interval. Next, two of the extrema of D_n are at a and b, since otherwise D_n would have at least $n+1$ interior extrema, each of which would be a zero of D_n'. (None of these points could be anything but a simple zero of D_n', since otherwise D_n' would have at least $n+2$ zeros (counting multiplicity), and $(D_n')^{(n)}=D_n^{(n+1)}=\phi^{(n+1)}$ would have a zero, contrary to hypothesis.) Then D_n' would have at least $n+1$ changes of sign, whereas we know that it has exactly n.

Now suppose, contrary to what we want to prove in Lemma 4, that $0<\phi^{(n+1)}(x) <f^{(n+1)}(x)$ on $[a, b]$ and $|f(x)-Q_n(x)|<E_n[\phi]$ on $[a, b]$, where Q_n is the polynomial (of degree at most n) of best approximation to f. Then at the points where $|\phi(x)-P_n(x)|=E_n[\phi]$, the function

$$F(x) = \phi(x) - P_n(x) - f(x) + Q_n(x) = D_n(x) - [f(x) - Q_n(x)]$$

has the sign of $\phi(x)-P_n(x)$. We already know that $D_n(x)$ has exactly $n+1$ changes of sign on (a, b); hence so does $F(x)$. Now

$$F^{(n+1)}(x) = \phi^{(n+1)}(x) - f^{(n+1)}(x) < 0$$

on $[a, b]$ by hypothesis, and so by Lemma 3 (applied to $g=-F$), $F(b)<0$. But $F(b)$ has the sign of $\phi(b)-P_n(b)=D_n(b)$, since b is a point where $|D_n(x)| =E_n[\phi]$. By Lemma 3 applied to $D_n(x)$, which has $D^{(n+1)}(x)=\phi^{(n+1)}(x)>0$, we must have $D_n(b)>0$, contradicting the facts that $F(b)<0$ and that $F(b)$ and $D_n(b)$ have the same sign.

This shows that we cannot have $E_n[f]<E_n[\phi]$; to complete the proof of Lemma 4, we need to know that $E_n[f]=E_n[\phi]$ is also impossible. Suppose the contrary, and let $0<\phi^{(n+1)}(x)<f^{(n+1)}(x)$ and $E_n[f]=E_n[\phi]$. Let λ be a positive number and consider $\phi(x)+\lambda f(x)=g(x)$. Since $g^{(n+1)}(x)=\phi^{(n+1)}(x)+\lambda f^{(n+1)}(x)$, we have $(1+\lambda)f^{(n+1)}(x)>g^{(n+1)}(x)>(1+\lambda)\phi^{(n+1)}(x)$, and hence by what has already been proved, $(1+\lambda)E_n[\phi]\leqq E_n[g]\leqq(1+\lambda)E_n[f]$. By assumption, $E_n[f]=E_n[\phi]$, so

$$E_n[g] = (1 + \lambda)E_n[\phi] = (1 + \lambda)E_n[f] = E_n[\phi] + E_n[f].$$

Let P_n and Q_n be the polynomials realizing $E_n[\phi]$ and $E_n[f]$; then

$$\begin{aligned} | P_n(x) + \lambda Q_n(x) - g(x) | &= | P_n(x) - \phi(x) + \lambda[Q_n(x) - f_n(x)] | \\ &\leqq E_n[\phi] + \lambda E_n[f]. \end{aligned}$$

But $E_n[g]=E_n[\phi]+E_n[f]$, so that no polynomial S_n of degree n (or less) can make $\max_x|S_n(x)-g(x)|$ less than this value. Since the best approximating polynomial for g is unique, it must therefore be $P_n(x)+\lambda Q_n(x)$, and consequently $|\phi(x)+\lambda f(x)-[P_n(x)+\lambda Q_n(x)]|$ attains its maximum value $E_n[\phi]+\lambda E_n[f]$ at $n+2$ points. This is possible only if $|\phi(x)-P_n(x)|$ and $|f(x)-Q_n(x)|$ attain their maximum values, namely $E_n[\phi]$ and $E_n[f]$, at the same $n+2$ points. This means that $F(x)=\phi(x)-P_n(x)-[f(x)-Q_n(x)]$ has $n+2$ double zeros, and so $F^{(n+1)}(x)$ has at least one zero; but we had $F^{(n+1)}(x)<0$ by hypothesis. This completes the proof of Lemma 4.

LEMMA 5 (Another theorem of Chebyshev). *On any interval of length* $2h$' $E_n[x^{n+1}]=2^n h^{n+1}$.

Another way of stating this is to say that if a polynomial of degree $n+1$ has its absolute value bounded by 1 on an interval of length $2h$, the absolute value of its leading coefficient is at most $2^{-n}h^{-n-1}$ (and can attain this value). This can be proved in a quite elementary way; see, for example, [**21**], p. 24, or [**7**].

We can now prove Lemma 1. Suppose that $f^{(n+1)}(x)>N>0$. Take $\phi(x) =Nx^{n+1}/(n+1)!$, so that $\phi^{(n)}(x)=N$; we then have $f^{(n+1)}(x)>\phi^{(n+1)}(x)>0$. By Lemma 4, $E_n[f]>E_n[\phi]$. By Lemma 5, $E_n[\phi]=2N(h/2)^{n+1}/(n+1)!$, and therefore $E_n[f]>2N(h/2)^{n+1}/(n+1)!$. If P_n is the polynomial of degree at most n that realizes $E_n[f]$, and R_n is any other polynomial of degree at most n, we have

$$E_n[f] = \max_x | f(x) - P_n(x) | \leqq \max_x | f(x) - R_n(x) |.$$

Take $R_n(x) \equiv 0$; then

$$2N(h/2)^{n+1}/(n+1)! < E_n[f] \leqq \max_x |f(x)|.$$

This establishes Lemma 1.

This article is based on a talk given at a seminar in honor of D. V. Widder, May 8, 1971.

References

1. S. Bernstein, Leçons sur les propriétés extrémales et la meilleure approximation des fonctions analytiques d'une variable réelle, Gauthier-Villars, Paris, 1926.
2. S. Bernstein, Sur les fonctions régulièrement monotones, Atti Congresso Int. Mat., Bologna, 2 (1930) 267–275.
3. R. P. Boas, Jr., Asymptotic relations for derivatives, Duke Math. J., 3 (1937) 637–646.
4. ———, Functions with positive derivatives, Duke Math. J., 8 (1941) 163–172.
5. ———, Representations for completely convex functions, Amer. J. Math., 81 (1959) 709–714.
6. ———, A Primer of real functions, Carus Mathematical Monographs, no. 13, 1960.
7. ———, Inequalities for the derivatives of polynomials, Math. Mag., 42 (1969) 165–174.
8. ——— and G. Pólya, Influence of the signs of the derivatives of a function on its analytic character, Duke Math. J., 9 (1942) 406–424.
9. ——— and D. V. Widder, Functions with positive differences, Duke Math. J., 7 (1940) 496–503.
10. P. J. Davis, Interpolation and approximation, Blaisdell, Waltham, Mass., 1963.
11. H. D. Kloosterman, Derivatives and finite differences, Duke Math. J., 17 (1950) 169–186.
12. A. Kolmogorov, On inequalities between the upper bounds of the successive derivatives of an arbitrary function on an infinite interval, Amer. Math. Soc. Translations, no. 4 (1949).
13. D. Leeming and A. Sharma, A Generalization of the class of completely convex functions, in Inequalities. III. Academic Press, New York, (to appear).
14. J. E. Littlewood, A Mathematician's miscellany, Methuen, London, 1953.
15. G. Pólya, On functions whose derivatives do not vanish on a given interval, Proc. Nat. Acad. Sci. U.S.A., 27 (1941) 216–217.
16. ———, On the zeros of the derivatives of a function and its analytic character, Bull. Amer. Math. Soc., 49 (1943) 178–191.
17. M. H. Protter, A generalization of completely convex functions, Duke Math. J., 24 (1957) 205–213.
18. A. C. Schaeffer, On the oscillation of differential transforms, III, Oscillations of the derivatives of a function, Trans. Amer. Math. Soc., 54 (1943) 278–285.
19. I. J. Schoenberg, On finite-rowed systems of linear inequalities in infinitely many variables, II, Trans. Amer. Math. Soc., 35 (1933) 452–478.
20. ——— and A. Cavaretta, Solution of Landau's problem concerning higher derivatives on the halfline, University of Wisconsin, MRC Tech. Rep. no. 1050 (1970).
21. J. Todd, Introduction to the constructive theory of functions, Academic Press, New York, 1963.
22. R. Welland, Bernstein's theorem for Banach spaces, Proc. Amer. Math. Soc., 19 (1968) 789–792.
23. D. V. Widder, Completely convex functions and Lidstone series, Trans. Amer. Math. Soc., 51 (1942) 387–398.
24. ———, The Laplace transform, Princeton University Press, 1941.

WALTER BURTON FORD

C. V. NEWSOM, Retired Vice-President, RCA

The death of Walter Burton Ford in Seneca County, New York, on February 24, 1971, when he was within a few months of his ninety-seventh birthday, represented a distinct break with the early history of the Association and, in fact, with the early history of the introduction of modern mathematics into the United States. Dr. Ford's educational background was typical of that of many of our distinguished mathematicians of the first part of the twentieth century. After graduating from the State Normal School located in his home community of Oneonta, New York, where he studied the curriculum, chiefly secondary school subjects, then required of prospective grade-school teachers, he registered at Amherst College in 1893. Since Ford had a strong interest in mathematics, stimulated originally by his efforts to understand the motion of the Great Comet of 1882, he decided after two years at Amherst to withdraw from the institution and enroll at Harvard: there he hoped to pursue an intensive program of studies in mathematics. At Harvard he received his A.B. degree, *magna cum laude*, in 1897 and his M.A. degree in 1898.

Ford then decided that the time had come to obtain a job, but he soon discovered that there was little demand for a person who had specialized in mathematics—a condition that has always existed in the United States except for the recent period after World War II. He finally became a teacher at the Albany Academy, an experience that he did not enjoy, so an offer in 1900 of an instructorship at the University of Michigan came to him, he later noted, "as if from heaven." The demand for additional instructors at Michigan was caused by the sudden influx of Spanish-American War veterans, and Ford in one of his memoirs commented that his "duties consisted of a heavy load of teaching freshmen only, and in classes so large that students were using radiators as well as chairs and benches for seats." "But," he observed, "at last I was started on my chosen career."

At Michigan, Ford worked closely with senior members of the staff, notably Alexander Ziwet, and he continued to maintain his contacts with the mathematicians at Harvard. His studies at that time led him to develop an intense interest in infinite series. In fact, he became convinced that mathematicians were only beginning to appreciate the significance of infinite series in mathematical analysis. He gave special attention to the use of series in the solution of certain types of differential equations, an interest that was awakened by his study of several papers of U. Dini, an Italian mathematician. Ford's first published paper, in the *AMS Bulletin* in 1901, was entitled, "Dini's Method for the Divergence of Fourier Series and other Allied Developments." Two years later, he had a short paper on Maclaurin's Series accepted for publication by the *Journal de Mathêmatique.*

After three years at Michigan, Ford decided, after consultation with the Harvard faculty, to spend a year in Europe, chiefly in France and Italy, so that

he could become familiar with some of the new mathematical ideas being promulgated by a brilliant generation of mathematicians in that part of the world. He especially anticipated the opportunity of meeting Dini, who spent many hours a day with Ford over a period of several weeks patiently discussing trends in mathematics.

When the year abroad was concluded, Ford returned home to a one-year appointment at Williams College. While at Williams, as a result of the inspiration gained from his European experiences, he prepared a paper entitled, "On the Problem of Analytic Extension as Applied to Functions Defined by Power Series." He was very proud of the work and submitted it to the Harvard faculty, then composed of such men as Bôcher, Osgood, and Byerly, to fulfill the final requirements for the doctorate. Unfortunately, the men at Harvard questioned the significance of the work. Ford immediately sent the paper to the French editors of *Journal de Mathêmatique*, who had published his earlier article and with whom he had become acquainted on his European trip. Shortly he received a very complimentary letter from the Frenchmen. Consequently, the Harvard faculty withdrew the earlier criticisms and awarded him the doctorate in 1905. The dissertation was published in 1906 in the *AMS Transactions*.

After receiving his doctorate, Ford rejoined the mathematics faculty of the University of Michigan where he stayed for the remainder of his academic career, having become a professor in 1917. During the year 1928–29, he lectured at universities in Holland, Belgium, France, and Italy under the auspices of the Carnegie Institute for International Peace. In 1940, he retired to the beautiful home that he had built during the 1930's overlooking Lake Cayuga in New York State; through the years that followed many mathematicians were his house guests, and at the time of the summer meeting at Cornell in 1946 two hundred members of the Association attended a lawn party at the Ford home.

Within a short time after receiving his doctorate, Ford became extremely interested in three long and difficult papers published during the years, 1906–1908, by E. W. Barnes, English mathematician who later became a distinguished but controversial Anglican bishop. Barnes employed in an ingenious manner the classical calculus of residues of the theory of functions of a complex variable to study the asymptotic behavior of functions defined by Taylor's Series. A far-reaching research program which Ford soon initiated, representing an extension of that of Barnes, became the dominant feature of his intellectual life from approximately 1910 until his death. Ford's research activities led to numerous publications, including articles in French and Italian journals and two monographs in the Scientific Series of the *University of Michigan Studies*, and they also provided the basis for a series of doctoral dissertations. Even in the last days of his life, in spite of failing eyesight, he was following his conviction that his methods, adapted from those of Barnes, would open the door to a proof of the hypothesis of Riemann pertaining to the zeros of his Zeta Function. An examination of the vast amount of work left on his desk after his death reveals the active state of Ford's mind even when he was in his middle nineties. As a foun-

dation for the proof which he was attempting he correctly derived a number of new formulas. But, alas, at the heart of his effort to prove the Riemann hypothesis he was duplicating an error committed several times in the past.

It is generally admitted that the creation of the Association with purposes essentially different from those of the Society resulted from the insight and determination of Herbert Ellsworth Slaught. But working behind the scenes during the first decade of the Association's existence were three men, intimate friends, who believed that a necessary Renaissance in American mathematics could be brought about only by the institution of major modifications in instructional materials and practices employed in American classrooms. The three men were the shrewd politically-oriented E. R. Hedrick, then of the University of Missouri, who was the first President of the Association in 1916, the aristocratic and stubborn R. C. Archibald of Brown University, who was President of the Association in 1922, and W. B. Ford, who was President of the Association in 1927 and was Editor of the MONTHLY from 1922 to 1927. Partly because of Hedrick's persuasion, Ford, in the late teens and early twenties, became co-author of an entire series of text-books for the secondary schools; the books, designed to represent a preliminary break-away from the Wentworth tradition, were widely used. In addition, Ford was the author of a beginning Calculus, designed to make a partial break with the Granville tradition; he also produced a very popular College Algebra and an elementary Analytic Geometry. Hedrick himself had already taken a major step toward the introduction in this country of the newer European thought in mathematics when he developed an English translation of Goursat's lectures at the University of Paris; the first volume appeared in 1904.

Ford, as all his students will recall, believed that principles which good mathematicians must follow in the development of any mathematical exposition should be consistent with accepted ideas pertaining to mathematical rigor. Moreover, he often made the statement, "I do not care how it is written, so long as it is clear." Osgood of Harvard was his ideal of the master expositor. Some students always left Ford's classes in dismay because they found it virtually impossible to explain mathematical concepts and principles in a way that the teacher would regard as acceptable. A doctoral candidate under his supervision could always expect to prepare at least twenty drafts of his dissertation before its linguistic format would be approved.

Ford's father, a prominent member of the business community of Oneonta, purchased stock during the period, 1905–1910, in the Computer-Tabulating-Recorder Company of Binghamton, New York. That company became the nucleus of a new company created in 1914 by the young supersalesman, Thomas J. Watson. Thus upon the death of his father in 1923, Ford, who seemed to possess his father's instinctive wisdom in the handling of investments, became one of the early stockholders in IBM. Few persons knew of Ford's wealth, for he shunned any publicity in connection with his many philanthropies. Ten or more colleges and universities profited handsomely from his great generosity,

and he was a contributor to a vast number of worthy causes. He erected a fine municipal building and a remarkable community library, a memorial to his deceased wife, in the Finger Lakes village of Ovid, New York. The Walter B. Ford Music Building, with its large pipe organ, on the campus of Ithaca College was named for him as a tribute to his generosity and to his life-long interest in music. During the twenties, Ford quietly took care of numerous small expenses of the young and struggling Association, and he was the largest contributor to the Chauvenet Fund, initiated by J. L. Coolidge in 1925. Ford's efforts in behalf of American mathematics and American society revealed the depth of his dedication to the concept of service.

THE NUMBER OF NUMERICAL OUTCOMES OF ITERATED POWERS

F. GÖBEL, Technological University Twente, and
R. P. NEDERPELT, Technological University
Eindhoven, The Netherlands

1. Introduction and summary. A notational convention, adopted in this paper, is to write the exponentiation of the real number a to the power b (commonly denoted by a^b) as $a \uparrow b$, just as is customary in some programming languages. Since exponentiation is not associative, expressions like $a \uparrow b \uparrow c$ are ambiguous, unless the order of application of the two operators under consideration is prescribed. We purposely dismiss the existing convention to interpret $a \uparrow b \uparrow c$ as $a \uparrow (b \uparrow c)$.

We focus our attention on expressions which are constructed by a proper insertion of a set of pairs of parentheses in a string of symbols of the form '$x \uparrow x \uparrow \cdots \uparrow x$', where x is a real number greater than 1, occurring $n+1$ times in the string. We call such a string of symbols, without inserted parentheses, an **iterated power.** Note that there is only one real number involved in the string.

The insertion of the pairs of parentheses must be such that the resulting expression can lead naturally and unambiguously to a **numerical outcome** for a given x and n. That is to say, the actual calculation, induced by the operators and the parentheses, has to deliver just one final result.

The string of symbols we get after the insertion of such a set of parentheses—with no calculation performed—is called a **bracketing of degree** n. When we admit total freedom in the way of inserting these parentheses, on the condition that we get a meaningful and unambiguous bracketing of degree n, we can obtain for given x and n a certain number of outcomes, which we denote by $L_n(x)$. In this paper we shall investigate this number.

A generally smaller number $N_n(x)$ of outcomes will be obtained when we admit only **nested** sets of parentheses, in the sense that each opening bracket precedes each closing bracket.

The number $N_n(2)$ was asked for in this MONTHLY as an Elementary Problem by G. Eldredge [1]. A solution given by M. Goldberg [2] was based on an incorrect proof, but one of us (Nederpelt) has adjusted Goldberg's proof, by applying a lemma of K. A. Post [4] in a slightly altered version.

As to $L_n(x)$, one may expect that $L_n(2)$ behaves differently from $L_n(x)$ for $x>1$ and $x\neq 2$, in view of the 'unique' relation $2\uparrow(2\uparrow 2)=(2\uparrow 2)\uparrow 2$. However, it is less obvious whether $L_n(3)$ and $L_n(5)$, say, are equal, or behave in comparable ways.

The number of different bracketings of degree n is equal to the Catalan number

$$\frac{1}{n+1}\binom{2n}{n},$$

as is shown, e.g., in [3, p. 25–26]. Hence,

$$L_n(x) \leqq \frac{1}{n+1}\binom{2n}{n} \tag{1}$$

for all n, and all $x>1$.

However, this upper bound for $L_n(x)$ is not sharp, even for $x\neq 2$, for different bracketings may yield the same outcome, as is shown by $(x\uparrow x)\uparrow(x\uparrow x)$ and $(x\uparrow(x\uparrow x))\uparrow x$. An upper bound which is much sharper than the one in (1) will be given for $x\neq 2$ in Section 2 and for $x=2$ in Section 3. In Section 4 we show that these upper bounds will also deviate from $L_n(x)$ for certain values of x. Section 5 states that the number of these exceptional values is countable.

2. A nontrivial upper bound for $L_n(x)$. We define a bracketing formally as a sequence of symbols which is either the sole symbol x or $(B_1\uparrow B_2)$, where B_1 and B_2 are bracketings. The **degree** of a bracketing is the number of symbols '$\uparrow$' occurring in it. Hence a bracketing of degree n contains n pairs of brackets, provided the final (outer) pair is not omitted.

We shall make a mapping from the set of bracketings of degree n onto the set of rooted trees with n edges. This mapping will be such that two bracketings of degree n, which are mapped on the same rooted tree, will have the same outcome for any given x. We shall construct this mapping as the composition of two mappings, described under the headings 'Standard Form' and 'Rooted Tree' respectively.

Standard Form. Let a bracketing of degree n be given. First we transform this bracketing into **standard form** by the following algorithm, applicable to the bracketing B:

1. Write the symbol 'log' in front of B.
2. If B is the sole symbol x, then proceed to step 4.
3. If B has the form '$(B_1\uparrow B_2)$', where B_1 and B_2 are bracketings, then replace 'log B' with '$B_2\cdot\log B_1$' and rename B_1 as B. Proceed to step 2.

4. An expression of the form 'prod $\cdot \log x$' has been obtained, where prod is a (possibly empty) formal product of bracketings.
   Replace 'prod $\cdot \log x$' with 'x' if prod is empty, with '($x \uparrow$ prod)' if prod consists of one factor, or with '($x \uparrow$ (prod))' if prod consists of two or more factors.
5. Apply the algorithm to each of the formal factors of 'prod', if any.

Bracketing B_1 in step 3 has a degree smaller than the degree of the bracketing B under consideration. So after a finite number of steps, we reach step 4. Each of the formal factors considered in step 4 is a bracketing of a degree smaller than the degree of the bracketing started with in the corresponding step 1. Besides, the number of formal factors is finite, so the algorithm terminates after a finite number of steps, as can be proved by induction.

If we read the formal product of bracketings, prod, as a product of expressions in a given x, we see that the algorithm does not change the outcome. In fact, nothing is done but to take logarithms, to reorder, and to exponentiate again. The use of the symbol 'log' only serves to make the algorithm clearer. It follows that the same outcome can be assigned to the standard form as to the original bracketing. The algorithm is unique, so the standard form is unique.

A more efficient algorithm may be given, but for our present purpose the one above is preferable.

(Let us define recursively a **natural bracketing** as either 'x' or '$(x \uparrow N)$', where N is a natural bracketing. The algorithm leaves unchanged these bracketings.)

Example. The bracketing

$$(((x \uparrow (x \uparrow x)) \uparrow x) \uparrow (x \uparrow x))$$

will be replaced successively by

$$\begin{aligned}&\log(((x \uparrow (x \uparrow x)) \uparrow x) \uparrow (x \uparrow x)),\\ &(x \uparrow x)\cdot\log((x \uparrow (x \uparrow x)) \uparrow x),\\ &(x \uparrow x)\cdot x\cdot\log(x \uparrow (x \uparrow x)),\\ &(x \uparrow x)\cdot x\cdot(x \uparrow x)\cdot\log x,\end{aligned}$$

and finally by

$$(x \uparrow ((x \uparrow x)\cdot x\cdot(x \uparrow x))).$$

In the last expression, the three formal factors of prod are $(x \uparrow x)$, x and $(x \uparrow x)$, respectively. They are left unchanged if the algorithm is applied to them, hence we may conclude that the standard form is obtained.

Rooted Tree. Next, a rooted tree is assigned to a standard form according to the following three rules:

1. Each x in the standard form corresponds to a node; the first x corresponds to the root.

2. If two x's are separated (apart from brackets) only by one '$\uparrow$', then the corresponding nodes are joined by an edge.
3. If an x is preceded immediately (apart from brackets) by the multiplication symbol '$\cdot$', then the node corresponding to this x is joined to the node corresponding to the last x preceding the first factor in the pertaining product.

Example. Consider the standard form

$$(a \uparrow ((b \uparrow c)\cdot(d \uparrow ((e \uparrow (f\cdot g))\cdot(h \uparrow j))))),$$

where, for ease of reference, the x's have been replaced with different letters. There are 9 letters, hence 9 nodes; the letter a corresponds to the root. According to the above rules, we get the tree shown in Figure 1.

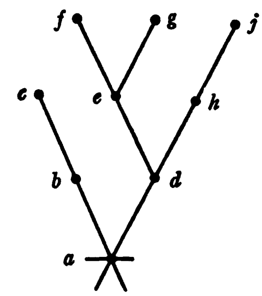


FIG. 1

THEOREM 2.1. *If a bracketing of degree n is transformed in the above way to a rooted tree, then this rooted tree has n edges.*

Proof. The number of x's in the bracketing is equal to the number of x's in the standard form, which is equal to the number of nodes in the tree. The degree of the bracketing is the number of x's minus one, and the number of edges of a tree is the number of nodes minus one. It follows that the degree of the bracketing is equal to the number of edges of the tree.

THEOREM 2.2. *An upper bound for $L_n(x)$ is given by the inequality $L_n(x) \leqq r_n$, where r_n is the number of rooted trees with n edges.*

Proof. If two bracketings lead to the same tree, then they have standard forms which can only differ in the order of their formal factors. Hence these bracketings yield the same numerical outcome, and, therefore, the number of outcomes is at most equal to the number of rooted trees.

It is known (see, e.g., [**5**, p. 127]) that r_n satisfies the following recursive relation:

$$r_n = \sum_{\psi(n)} \prod_i \binom{r_{i-1} + m_i - 1}{m_i}, \tag{2}$$

where the sum extends over all **partitions** $\psi(n)=1^{m_1}2^{m_2}\cdots$ of n, i.e., over all ordered sets of integers $m_1, m_2, \cdots$ with $\sum j\, m_j = n$, where $m_j \geqq 0$.

Some values of $L_n(4)$ and r_n are given in the following table:

n	1	2	3	4	5	6
$L_n(4)$	1	2	4	9	20	48
r_n	1	2	4	9	20	48

It is quite tempting to conjecture that $L_n(x)=r_n$ for all n and all $x>1$, with $x=2$ excluded.

3. An upper bound for $L_n(2)$. When $x=2$, we can improve considerably on the upper bound r_n. This is done by making use of the fact that $x \uparrow x$ and $x\cdot x$, which can both be parts of standard forms, give the same outcome for $x=2$.

Let the **order** of a node (of a tree) be defined as the number of edges which have that node as one of their end-points. By a **terminal edge** of a rooted tree we understand an edge which has as one of its end-points a node different from the root, with order 1. A **trimmed tree** is a rooted tree in which each terminal edge has as one of its end-points a node of order $\geqq 3$.

Let s_n be the number of trimmed trees with n edges.

THEOREM 3.1. *An upper bound for $L_n(2)$ is given by the inequality $L_n(2) \leqq s_n$.*

Proof. Let a bracketing of $2 \uparrow 2 \uparrow \cdots \uparrow 2$ be given. After bringing it into standard form, we replace each sequence of the form '$(2 \uparrow 2)$' with '$2\cdot 2$' if '$(2 \uparrow 2)$' is a factor in a product of two or more factors, or else with '$(2\cdot 2)$'. To this modified standard form, which has the same numerical outcome (and the same number of 2's) as the original one, we assign a rooted tree in the same way as in Section 2. It is clear that the result is a trimmed tree, and the theorem follows.

THEOREM 3.2. *Let c_n be defined by: $c_1=1$, $c_2=0$, and $c_n=s_{n-1}$ for $n\geqq 3$. The number of trimmed trees with n edges is given by*

$$s_n = \sum_{\psi(n)} \prod_i \binom{c_i + m_i - 1}{m_i}, \tag{3}$$

where the sum extends over all partitions $\psi(n)=1^{m_1}2^{m_2}\cdots$ of n.

Proof. The quantity c_n can be interpreted as the number of trimmed trees with n edges and for which the order of the root ('the number of stems') is 1. The reasoning which leads to (2) can be duplicated in this case, and leads to (3).

Some values of s_n are given below.

n	1	2	3	4	5	6	7	8	9	10
s_n	1	1	2	4	8	17	36	79	175	395

4. Some complications. The conjecture at the end of Section 2 is false, as is shown by the following counterexample.

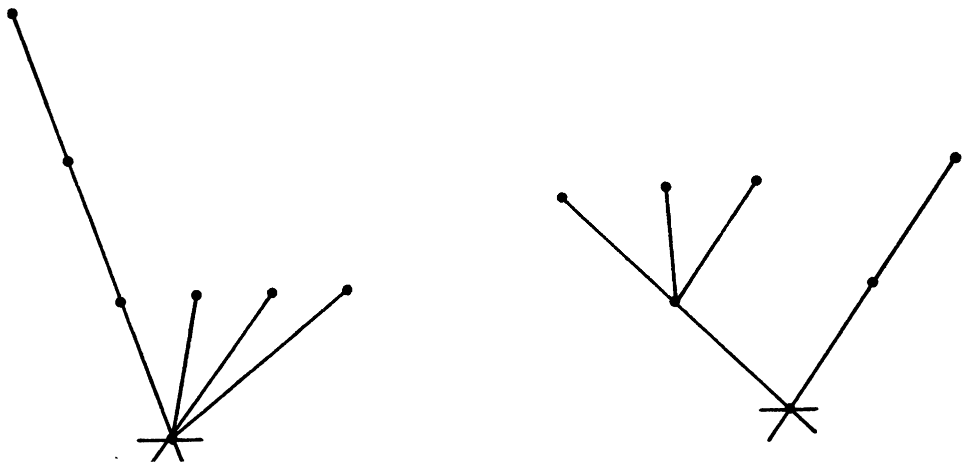


FIG. 2

Let $n=6$ and $x=3$. The two rooted trees depicted in Figure 2 lead to the same numerical outcomes.

When $x=2$, the two trimmed trees of Figure 3 also lead to the same numerical result. Hence, from $n=8$ onward, s_n too is only an upper bound.

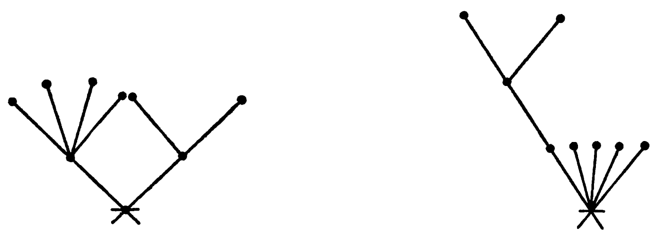


FIG. 3

If we look deeper into these examples, we see that in both cases two mutually different subtrees, for which the numerical outcomes are the same for the respective values of x, are interchanged.

For any integer value of x, counterexamples of this type can be constructed.

If $L_n(x) < r_n$ for some x_0 and some n_0, then, of course, the inequality holds for x_0 and all $n \geqq n_0$. Hence, for each integer x there exists an n_0 such that $L_n(x) < r_n$ for $n \geqq n_0$.

It is easy to find a noninteger x with $L_n(x) < r_n$. Assume $n=3$. Then the bracketings $(((x \uparrow x) \uparrow x) \uparrow x)$ and $(x \uparrow ((x \uparrow x) \uparrow x))$ yield different standard forms: $(x \uparrow (x \cdot x \cdot x))$ and $(x \uparrow (x \uparrow (x \cdot x)))$, respectively, leading to different rooted trees. However, the outcomes for $x = \sqrt{3}$ are the same, hence $L_n(\sqrt{3}) < r_n$ for $n \geqq 3$.

5. The validity of the equality $L_n(x) = r_n$. Let n be given. Denote the standard forms of degree n by $S_1, \cdots, S_{r_n}$, and denote the corresponding numerical outcomes by $S_1(x), \cdots, S_{r_n}(x)$, as functions of x.

LEMMA. *If* $1 \leqq j < k \leqq r_n$, *then there exists an* $x > 1$ *such that* $S_j(x) \neq S_k(x)$.

The proof, which will be omitted, is based on the asymptotic behavior of $S_i(x)$ as $x \to \infty$, which implies the existence of a positive number N_{jk} such that $S_j(x) \neq S_k(x)$ for all $x > N_{jk}$.

THEOREM 5.1. *For each* n, *there is at most a countable number of* $x > 1$ *such that* $L_n(x) \neq r_n$.

Proof. If $f(z)$, $g(z)$, and $h(z)$ are analytic functions in the open right half plane H, then so are $z \uparrow f(z)$ and $g(z) \cdot h(z)$. Because of this property, $S_j(z)$ is analytic in H, as a look at the structure of the standard form S_j shows. Here $S_j(z)$ is the complex function of the complex variable z, identical to $S_j(x)$ for real $z = x$.

In each closed, simply-connected subset of H, equality of $S_j(z)$ and $S_k(z)$, where $k \neq j$, can occur for at most a finite number of z, as a consequence of the lemma. Since H can be covered by a countable number of closed, simply-connected subsets, the complex equation $S_j(z) = S_k(z)$, where $j \neq k$, has at most a countable number of solutions in H.

Hence, the same is true for the real equation $S_j(x) = S_k(x)$, where $j \neq k$, and $x > 1$. If x is not a solution of the last equation for some j and k, where $1 \leqq j < k \leqq r_n$, then there is a one-to-one correspondence between numerical outcomes and standard forms, and the theorem follows.

References

1. G. Eldredge, Problem E 1903, this MONTHLY, 73 (1966) 666.
2. M. Goldberg, Solution of problem E 1903, this MONTHLY, 77 (1970) 525.
3. M. Hall, Combinatorial Theory, Ginn-Blaisdell, Waltham, Mass., 1967.
4. K. A. Post, A combinatorial lemma involving a divergence criterion for series of positive terms, this MONTHLY, 77 (1970) 1085–87.
5. J. Riordan, An Introduction to Combinatorial Analysis, Wiley, New York, 1958.

ADDENDUM TO "THE USES OF ANALYTICITY IN OPERATOR THEORY"

ANGUS E. TAYLOR, University of California

I wish to repair two omissions in my article, *Notes on the History of the Uses of Analyticity in Operator Theory*, in the April 1971 issue of this MONTHLY. I should have mentioned the important contribution made by E. R. Lorch in his paper, *The Spectrum of Linear Transformations*, Trans. Amer. Math. Soc., 52 (1942) 238–248. Lorch shows how to use the Cauchy integral formula to obtain projections associated with spectral sets of an operator, and develops the associated reducibility theory fully. My failure to mention this paper is the more inexplicable since I was familiar with it and had referred to it in my own 1943 paper. Another oversight was my failure to include reference to the valuable book of E. Hille and R. S. Phillips, *Functional Analysis and Semi-Groups* (first edition 1948, revised edition 1957) published by the American Mathematical Society. It contains many results on analyticity in connection with resolvents and spectra, some of them published there for the first time. See particularly Chapters IV and V (1957 edition) for general theory. Actually, the whole book is full of complex variable methods in functional analysis, with numerous applications to particular operators and functional transformations.

MATHEMATICAL NOTES

EDITED BY DAVID DRASIN

Manuscripts for this Department should be sent to Robert Gilmer, Department of Mathematics, Florida State University, Tallahassee, FL 32306. Notes are usually limited to three printed pages.

COMPLEMENTS AND COMMENTS

DAVID DRASIN and ROBERT GILMER

The past year saw a marked decrease in the number of supplementary comments our readers have sent us. We are again very grateful to those who have passed the following information:

Calculus. Our most arresting comment comes from D. V. Anderson, who observes that the method used by M. R. Spiegel (1956, p. 35) to evaluate $\int_0^\infty e^{-x^2}dx$ by single variable methods was anticipated in 1842 by Augustus De Morgan on page 294 of his textbook.

Set theory and logic. J. Frederick Leetch writes that the proof of the Schroeder-Bernstein Theorem which appeared in 1968, p. 508, is much like that given by M. Reichaw-Reichbach, Colloquium Mathematicum (1955, p. 163).

Geometry. In the April 1971 MONTHLY, pp. 384–5, M. V. Subbarao sought to determine the number of triangles $N(\lambda)$ whose integer-valued sides add up to

λ times their area. Three readers have pointed out a gap in Subbarao's proofs: R. Jones, M. J. Marsden, and D. L. Shell. In particular, Subbarao had claimed that $N(\lambda)$ is zero for all λ greater than $8^{1/2}$, except the $12^{1/2}$. However the equilateral triangle, each side of which is of unit length, works for $\lambda = 48^{1/2}$.

Algebra. Josef Schmid showed (1970, p. 998) that if A and B are $n \times m$ and $m \times n$ matrices over a unitary commutative ring R and I_n and I_m denote the unit matrices of orders n and m, then

$$t^m \det(tI_n - AB) = t^n \det(tI_m - BA).$$

This result was anticipated in an article by J. L. Brenner, pp. 531–2 which appeared in the October 1968 issue of Linear Algebra and Its Applications.

K. K. Butler has pointed out an item of possible confusion in R. Korfhage's article "Solutions of $X^2 = I$ for matrices over finite rings with unity" (1968, pp. 634–6). It is important to assume that $1+1$ is not equal to 1; without this assumption Butler has counterexamples to Korfhage's results. Korfhage suggests that perhaps he should have used the term "Boolean ring" rather than "Boolean algebra."

In 1970, pp. 743–5, G. J. Simmons deduced formulas for the number of irreducible polynomials of degree n over $GF(p)$. However, the method used is also suggested in exercises which appear in two standard textbooks: volume 3 of Jacobson's *Lectures in Abstract Algebra* (ex. 1, p. 61) and in Hasse's *Exercises to Higher Algebra* (ex. 18, p. 186). Further, Theorem 3 of Simmons' paper and its proof appears as Theorem 308, p. 497 of *Algebra*, Part 1, by L. Redei.

R. Kopperman has pointed out that for any field F, the semigroup ring of Q_0, the additive semigroup of non-negative rationals, over F is an example of a non-Noetherian ring with identity which is easily accessible to beginners. Other examples are given by R. Gilmer on pp. 621–623 (1970).

R. L. Roth, on pp. 392–3 of this volume, discussed algebraic extensions of the rationals by square roots of successive primes. This article brought several comments from D. J. McCarthy. He first observes that from the theorem of the primitive element, we know that $Q(\sqrt{p_1}, \cdots, \sqrt{p_n}) = Q(\theta)$ for some θ. It is not difficult to show that in fact one can take $\theta = \sqrt{p_1} + \cdots + \sqrt{p_n}$; this can be seen by examining the classical demonstration of the existence of a primitive element as in van der Waerden (vol. I, 1953, p. 126). In order to apply this procedure here we need that $\sqrt{p_n} \notin Q(\sqrt{p_1}, \cdots, \sqrt{p_{n-1}})$. It follows that this particular element θ has degree 2^n over Q. (An interesting challenge to students who are skeptical of the strength of this result is to seek a direct proof of the simpler fact that $\sqrt{p_1} + \cdots + \sqrt{p_n}$ is irrational.) In seeking an example of an algebraic extension of infinite degree over Q which is "easily grasped intuitively," there are simple alternatives to the one suggested by Roth. For example, it is clear that if p is a prime and $n > 1$, then $X^n - p$ is irreducible over Q (e.g., via Eisenstein). Thus $[Q(\sqrt[n]{p}:Q] = n$. Since $Q(\sqrt[n]{p})$ is a subfield of $E = Q(\sqrt{p}, \sqrt[3]{p}, \cdots, \sqrt[n]{p}, \cdots)$ for all n, it follows immediately that $[E:Q]$ is infinite. Enzo R. Gentile has also written to us about $Q\ (\sqrt{p_1}, \sqrt{p_2}, \cdots)$.

I. Kleiner has observed that the set $\{\sqrt[m_i]{k}\}$ (where $\{m_1, m_2, \cdots\}$ is a sequence of positive integers) is linearly independent over Q if k is a square-free positive integer and $(m_i, m_j)=1$ for $i\neq j$. Hence the field $Q(\{\sqrt[m_i]{k}\})$ is an infinite algebraic extension of Q.

Finally, all these results can be generalized if square roots are replaced by kth roots. This result (attributed to I. J. Richards) may be found in Gaal, *Classical Galois Theory* (1971), pp. 234–237.

Number theory. J. L. Paul presented a derivation of $\sum_{i=1}^{n} i^k$ (1971, pp. 271–2). It is possible to give a simple induction proof of this result as J. Behboodian, F. B. Correia, and S. L. Gupta have noted. Paul's article presents an alternative and very elegant way of obtaining the same result.

In the 1970 MONTHLY, pp. 848–852, an article by S. W. Golomb appeared on powerful numbers. A positive integer is powerful if, whenever it is divisible by a prime number p, it is also divisible by p^2. R. Stanley has shown, and will publish elsewhere, several results which resolve problems mentioned in this paper. In particular, he has shown that every nonzero integer can be represented as a difference of two powerful numbers in infinitely many ways. There are also infinitely many consecutive integers, both powerful and neither squares.

Stanley's methods use algebraic number theory. However, A. Makowski has presented an elementary proof that infinitely many positive integers are so representable and this will appear in the Mathematical Notes section sometime next year.

Analysis. Several comments should be made about S. G. Wayment's article on continuity-differentiability relationships (1970, pp. 740–3). In 1951, p. 408, M. K. Fort proved that a real function, which is discontinuous on a dense set, can only be differentiable on a set of first category. A now standard reference for this is Boas's *A Primer of Real Functions*, p. 126. Wayment claimed that Boas's proof and Fort's theorem were incorrect, but now withdraws the latter claim. Wayment had claimed in Section 4 that the maximum number of rationals with denominator g in (x, y) is $g(y-x)$. This is only true, as E. Stevenson has observed, if $g(y-x)$ is replaced by $g(y-x)+1$, and the set of Liouville numbers shows that the error is not removable. It is easy to fix up Boas's proof by using the ideas in the closing line of the first paragraph of Section 4 in Wayment's paper, since if f is discontinuous, then either the left limit or the right limit fails to equal f on a dense set. In any case, Fort's original proof is correct. Further discussions of these ideas will appear in a forthcoming article by E. M. Beesley, A. P. Morse and D. C. Pfaff.

D. Stone reports that an article by P. Erdös and J. C. Oxtoby (Trans. A.M.S., 79(1955), 91–102) contains a characterization of plane sets whose intersections with all measurable rectangles of positive measure have positive measure. Darst and Goffman rediscovered the sufficiency part of this in their note "A Borel set which contains no rectangles" (1970, pp. 728–9).

SOME ASSOCIATIVITY CONDITIONS FOR ALGEBRAS

RAYMOND COUGHLIN AND MICHAEL RICH, Temple University

A **non-associative algebra** is an algebra in which the associative law foı multiplication is not necessarily assumed. The study of such algebras does not usually involve doing away with all forms of associativity; one usually replaces the associative law with some weaker identity. The classical non-associative algebras utilize the associator $(x, y, z)=(xy)z-x(yz)$ to this end by requiring that the elements of the algebra satisfy an identity in which all associators of a particular form are zero. For instance, a **flexible algebra** satisfies $(x, y, x)=0$, an **alternative algebra** satisfies $(x, x, y)=(y, x, x)=0$, and a **Jordan algebra** is a commutative algebra satisfying $(x, y, x^2)=0$.

The question arises: What if, instead of placing restrictions on the entries, we assume that the set of all associators lies in some restricted subset of the algebra? As a partial answer to this question we provide elementary proofs to the following theorems:

THEOREM 1. *If A is a non-associative algebra with an identity element* 1 *over a field F and if $(x, y, z)\in F1$ for all x, y, z in A, then A is associative.* (By $F1$ we mean the set of elements of A which are scalar multiples of the identity of A.)

Proof. It can be verified directly that the Teichmuller identity

$$a(x, y, z) + (a, x, y)z = (ax, y, z) - (a, xy, z) + (a, x, yz) \tag{1}$$

holds in any algebra. Assume, then, that a, x, y, z are any elements of A and that $(x, y, z)\neq 0$. Then, since (a, x, y) and $(ax, y, z)-(a, xy, z)+(a, x, yz)$ are elements of $F1$, it follows that for every a in A, $a=\alpha(a)z+\beta(a)1$, where $\alpha(a)$ and $\beta(a)$ are in F. Therefore A is generated by the elements 1 and z. Also $x=\alpha(x)z+\beta(x)1$ and $y=\alpha(y)z+\beta(y)1$. Thus

$$(x, y, z) = (\alpha(x)z + \beta(x)1, \alpha(y)z + \beta(y)1, z).$$

Since the associator is linear in each of its variables, $(x, y, z)=\alpha(x)\alpha(y)(z, z, z)$. But $z^2=\alpha z+\beta 1$ for $\alpha, \beta\in F$. Therefore $z^2z=zz^2$ and $(z, z, z)=0$. Hence $(x, y, z)=0$, and A is associative.

The **center** C of an algebra A is defined by

$$C = \{c \in A \mid cx = xc \text{ and } (c, x, y) = (x, c, y) = (x, y, c) = 0 \quad \text{for all } x, y \in A\}.$$

THEOREM 2. *If A is a simple algebra and (x, y, z) is in C, the center of A, for all x, y, z in A, then A is associative.*

Proof. Since A is simple, either $C=0$ or C is a field [1, p. 291]. If $C=0$ we are finished. Otherwise A can be thought of as an algebra over C and the argument of Theorem 1 applies to show that A is associative.

Reference

1. N. Jacobson, Lie Algebras, Interscience Publishers, New York and London, 1962.

ON GENERALIZED FIBONACCI NUMBERS

M. D. MILLER (student), University of California at Los Angeles

In studying generalized Fibonacci numbers of order $k>1$ defined by the conditions $F^k(r)=0$ for $0\leqq r<k$, $F^k(k)=1$, and $F^k(n)=\sum_{i=1}^k F^k(n-i)$ for $n>k$, it is necessary to consider the polynomial equation $x^k-x^{k-1}-\cdots-x-1=0$, which we shall call the **characteristic equation** of the kth order Fibonacci sequence.

E. P. Miles, Jr., in this Monthly 67 (1960) 745–752, showed that the roots of this equation are distinct, and in addition, one root lies between 1 and 2, and the remaining $k-1$ lie within the unit circle of the complex plane. He showed this by reducing the equation to a form where Rouché's theorem could be applied.

We offer here a proof of the above assertions involving nothing more than elementary theory of equations.

THEOREM. *Let* $f(z)=z^k-z^{k-1}-\cdots-z-1$, *for* $k>1$. *Then* (a) *f has a real zero z_0 such that* $1<z_0<2$; (b) *the remaining $k-1$ zeros of f lie within the unit circle in the complex plane*; (c) *the zeros of f are simple.*

Proof. Let $g(z)=(z-1)f(z)=z^{k+1}-2z^k+1$. By Descartes's rule of signs, $f(z)$ has exactly one positive real zero, say $z=z_0$. Since $f(1)<0$ and $f(2)=1$, we see that $1<z_0<2$.

(1) For real $x>z_0$, we have $f(x)>0$, whereas $f(x)<0$ if $0<x<z_0$.
(2) For real $x>z_0$, also $g(x)>0$, whereas $g(x)<0$ if $1<x<z_0$.

We observe that $f(z)$ has no complex zero z_1 with $|z_1|>z_0$. For if $f(z_1)=0$, then

$$z_1^k = z_1^{k-1} + \cdots + z_1 + 1$$

and

$$|z_1|^k = |z_1^{k-1} + \cdots + z_1 + 1| \leqq |z_1^{k-1}| + \cdots + |z_1| + 1.$$

But this says that $f(|z_1|)\leqq 0$, which contradicts (1) above.

Furthermore, $f(z)$ has no complex zero z_2 with $1<|z_2|<z_0$. For if z_2 were such a zero, then $f(z_2)=g(z_2)=0$ and $2z_2^k=z_2^{k+1}+1$. From this we have $|2z_2^k| =|z_2^{k+1}+1|\leqq|z_2^{k+1}|+1$. This says that $g(|z_2|)\geqq 0$, contradicting (2) above.

We claim now that $f(z)$ has no complex zero $z_3\neq z_0$ with either $|z_3|=z_0$ or $|z_3|=1$. For if it had, then $g(z_3)=0$ and $2z_3^k=z_3^{k+1}+1$. Then

$$2|z_3|^k = |z_3^{k+1} + 1| \leqq |z_3|^{k+1} + 1.$$

Equality can hold here only if z_3^{k+1} is real, which forces z_3^k and also z_3 to be real since $g(z_3)=0$. By applying Descartes's rule to $g(z)$, we see that $f(z)$ has exactly one positive zero and either one or no negative zero, depending whether k is even

or odd. If k is even, we have $f(0) = -1$ and $f(-1) = 1$; thus no zero of $f(z)$, besides z_0, has absolute value 1 or z_0. The claim is proved.

To prove that $f(z)$ has simple zeros, suppose that this were not the case. Then also $g(z)$ would have a multiple zero, so $g(z)$ and $g'(z)$ would have a common zero. But the zeros of $g'(z)$ are $z=0$ and $z=2k/(k+1)$, whereas the only rational root of $g(z)=0$ is $z=1$.

ON CERTAIN DIOPHANTINE EQUATIONS

D. A. BUTTER, Joslyn Mfg. Co., Cleveland, Ohio (Now at ITT Lamp Div., Lynn, Massachusetts)

THEOREM. *If $x_1^p + \cdots + x_n^p = z^p$, where the x_i and z are positive integers, $n \geqq 2$, and p is an odd prime, then $x_1 + \cdots + x_n \geqq z + 2p$ and $p < \frac{1}{2}(n-1)z$.*

Proof. By Fermat's theorem, $x^p \equiv x \pmod{p}$. Also $x^p - x$ is even, so $x^p \equiv x \pmod{2p}$. Therefore $x_1 + \cdots + x_n \equiv z \pmod{2p}$. But $z^p = x_1^p + \cdots + x_n^p < (x_1 + \cdots + x_n)^p$, hence $x_1 + \cdots + x_n > z$, $x_1 + \cdots + x_n \geqq z + 2p$. Also $x_i < z$, so $nz > z + 2p$.

COROLLARY. *If p is an odd prime and $x^p + y^p = z^p$ for positive integers x, y, and z, then $p < \frac{1}{2}\min(x, y)$.*

Proof. Assume $x < y$. Then $x + y \geqq z + 2p$, hence $x - 2p \geqq z - y > 0$.

ANOTHER TOPOLOGICAL EQUIVALENT OF THE AXIOM OF CHOICE

S. P. FRANKLIN AND B. V. S. THOMAS, Carnegie-Mellon University

A well-known and sometimes useful consequence of Zorn's Lemma is that *any perfect onto map has an irreducible restriction to a closed subspace of its domain.* (Recall that a continuous function $f: X \to Y$ is called **perfect,** or sometimes **proper** [1], if the image of a closed set is closed, and if the preimage of each point of Y is compact. An onto function is **irreducible** if the image of any proper closed subset of the domain is a proper subset of the range.)

It is easily proved, but apparently less well known, that this assertion, even if restricted to T_1 spaces, is actually equivalent to the axiom of choice. Indeed, suppose $\{S_\alpha\}_{\alpha \in Y}$ is a non-empty pairwise disjoint family of non-empty sets. For each α let X_α be the set S_α provided with the cofinite topology (the complements of singleton sets form a subbase), and let X be their free union [2]. If the index set Y is given the discrete topology, the function $f: X \to Y$ which maps each X_α constantly to α is a perfect onto map since X_α is compact. The resulting subspace A of X on which f is irreducible must intersect each X_α in a singleton (if some $X_\alpha \cap A$ contained two points, A could be reduced by an open X_α-neighborhood of one of them) and provides, therefore, a choice function on $\{X_\alpha\}_{\alpha \in Y}$.

Gleason uses the existence of irreducible restrictions to prove that extremally

disconnected compact Hausdorff spaces are projectives [3]. The observation of the previous paragraph dims hopes of an "effective" construction of the factorization involved.

It is tempting to try to prove that the existence of irreducible restrictions implies the axiom of choice by taking each X_α to be some one point compactification, hence a Hausdorff space. It may have been this which led the referee to ask if existence of irreducible restrictions of perfect onto maps out of *Hausdorff* spaces also implies the axiom of choice. We leave this question to others, remarking only that, while the Tychonoff theorem is equivalent to the axiom of choice [5], the Tychonoff theorem for Hausdorff spaces is equivalent to the prime ideal theorem for Boolean algebras [6], a strictly weaker assertion [4].

References

1. N. Bourbaki, Topologie Générale, 3rd ed., Act. Sci. et Ind. No. 1142, Hermann, Paris, 1961, §10.
2. J. Dugundji, Topology, Boston, 1966, p. 127.
3. A. M. Gleason, Projective topological spaces, Illinois J. of Math., 2 (1958) 482–489.
4. J. D. Halpern, The independence of the axiom of choice from the Boolean prime ideal theorem, Fund. Math., 55 (1964) 57–66.
5. J. L. Kelley, The Tychonoff product theorem implies the axiom of choice, Fund. Math., 37 (1950) 75–76.
6. H. Rubin and D. Scott, Preliminary Report, Some topological theorems equivalent to the Boolean prime ideal theorem, Bull AMS, 60 (1954) 389.

THE DERIVATIVE OF THE TOTAL VARIATION FUNCTION

G. A. HEUER, Concordia College

Let f be a function of bounded variation on $[a, b]$ and for x in $[a, b]$, let $v_f(x)$ be the total variation of f on $[a, x]$. One of the fundamental properties of v_f which the student of analysis learns is that it is continuous at x if and only if f is continuous there. He may perhaps learn also that the set of points where f (and v_f) is discontinuous is countable, and possibly even that f and v_f are both differentiable except at a set of measure 0. It may occur to the curious student to ask whether the sets of points, where the derivative exists, are identical for f and v_f. The example $f(x) = |x|$ on $[-1, 1]$ shows this to be false, but suggests that the proper question might substitute "one-sided derivative" for "derivative." Of course the continuity property is significant for the subsequent theory while the question here raised is not. Nevertheless the question seems intrinsically interesting, and leads to some interesting examples. These might be suitable for student presentation at an undergraduate seminar; a more able student might be asked to obtain them independently as an honors project. Some such results and examples follow.

We consider right hand derivatives only. Note first that if f is monotone in some closed right hand (henceforth r.h.) neighborhood of c, one of $v_f - f$ and $v_f + f$ is constant on this neighborhood, so that v_f has a r.h. derivative if and

only if f has, and $v_f'(c) = |f'(c)|$. In the remaining case f oscillates in every r.h. neighborhood of c.

EXAMPLE 1. $f(x) = x^2 \cos(1/x)$, if $0 < x$; $f(0) = 0$. Then $f'(0) = 0$. To estimate $v_f(x)$, note that on each interval $[2/(2n+1)\pi,\ 2/(2n-1)\pi]$, f vanishes at the endpoints and has a single extreme point x_n, where $1/n\pi < x_n < 2/(2n-1)\pi$. Thus the variation of f on this interval is $2|f(x_n)|$, and

$$(1/n\pi)^2 = |f(1/n\pi)| < |f(x_n)| < x_n^2 < [2/(2n-1)\pi]^2.$$

By the integral test [1; p. 393],

$$1/n = \int_n^\infty dx/x^2 < \sum_{k=n}^\infty 1/k^2 < (\pi^2/2)v_f(2/(2n-1)\pi) < \sum_{k=n}^\infty [2/(2k-1)]^2$$
$$< \int_{(2n-3)/2}^\infty dx/x^2 = 2/(2n-3).$$

Then for $2/(2n-1)\pi \leqq x \leqq 2/(2n-3)\pi$ (with $n \geqq 3$), we have $1/n < (\pi^2/2)v_f(x) < 2/(2n-5)$, and hence $(2n-3)/n\pi < (1/x)v_f(x) < (4n-2)/(2n-5)\pi$. It follows that $v_f'(0) = 2/\pi$.

EXAMPLE 2. $f(x) = x^r \cos(1/x)$, if $0 < x$; $f(0) = 0$. For all $r > 1$, $f'(0) = 0$. By the obvious modifications of the calculation in Example 1 one finds that $v_f'(0) = 0$ if $r > 2$ and $v_f'(0) = \infty$ if $r < 2$.

The examples so far leave some hope that if infinite derivatives are admitted, v_f has a r.h. derivative at c if and only if f has. The facts, however, are something else.

THEOREM 1. *If $0 \leqq p \leqq q \leqq \infty$, there is a function f, differentiable and of bounded variation on $[0, 1]$, for which the lower and upper r.h. derivatives of v_f at 0 are, resp., p and q.*

Proof. Note first that given $K > 0$, $\epsilon > 0$, and $a < b$, we may construct a differentiable function f on $[a, b]$ such that $|f(x)| \leqq \epsilon$ for all x and the total variation of f on $[a, b]$ is K. (Moreover, we may do so even if f and f' are to have preassigned values at a and b, subject to $|f(a)|$ and $|f(b)| \leqq \epsilon$.) For the following construction we assume $0 < p < q < \infty$; the cases $p = 0$ and $q = \infty$ require easy modification; and the case $p = q$ is trivial. We form by induction a sequence $\{P_i\} = \{(x_i, y_i)\}$ of points with P_{2n} on the line $y = px$ and P_{2n-1} on the line $y = qx$, and such that $\{x_n\}$ and $\{y_n\}$ both decrease monotonically to 0, and $x_1 = 1$. We now construct f so that $|f(x)| \leqq x^2$ (and hence $f'(0) = 0$) and v_f passes through each P_i. On the interval $[x_{n+1}, x_n]$, choose f to have total variation $y_n - y_{n+1}$, with $|f(x)| \leqq x_{n+1}^2$ and $f(x_{n+1}) = 0$. For $n > 1$ we may also require that the left hand derivative of f at x_n equal the right hand one. Then $v_f(x_n) = y_n$, as required. To keep the graph of v_f between the lines $y = px$ and $y = qx$ (so that p and q are indeed the lower and upper r.h. derivatives of v_f at 0) it is sufficient

to keep $|f'(x)|$ constant in each interval $[x_{n+1}, x_n]$ except in a small neighborhood of each extreme point, where f may be smoothed enough for it to remain differentiable.

THEOREM 2. *If* $-\infty \leqq p \leqq q \leqq \infty$, *there is a function* f *of bounded variation on* $[0, 1]$ *for which the lower and upper r.h. derivatives at* 0 *are, resp.,* p *and* q, *and for which* v_f *has a r.h. derivative at* 0. *When both* p *and* q *are finite,* f *may be chosen so that* $v_f'(0)$ *is finite.*

Proof. Assume first that $-\infty < p < q < \infty$. The graph of f will consist of a sequence of line segments each joining a point of the line L_1: $y = px$ to one on the line L_2: $y = qx$. Choose $m > \max\{|p|, |q|\}$. Connect the point $P_1 = (1, q)$ to a point P_2 on L_1 by a segment of slope m; join P_2 to P_3 on L_2 by a segment of slope $-m$; then P_3 to P_4 on L_1 by a segment of slope m, etc. Clearly f has the required properties. Since on any interval $(a, 1)$ with $0 < a$ the variation of f is $m(1-a)$, and v_f is continuous at 0, we have $v_f(x) = mx$; thus $v_f'(0) = m$.

For the remaining cases, a modification of this idea works. For $q = \infty$, e.g., one replaces L_2 by a sequence $\{L_{2n}\}$ of lines whose slopes increase without bound. The only difficulty is to keep the variation of f bounded. This may be accomplished by choosing the points $P_{2n-1} = (x_{2n-1}, y_{2n-1})$ on L_{2n} so that $y_{2n+1} = (1/2)$ y_{2n-1} (for instance), and choosing the slopes of the connecting segments accordingly. The details will be left for the reader.

Reference

1. R. G. Bartle, The Elements of Real Analysis, Wiley, New York, 1964.

ON SEPARATION BY SPHERICAL SURFACES

S. R. LAY, University of California at Los Angeles, and Aurora College

In 1903 Paul Kirchberger [3] proved that two finite subsets P and Q of Euclidean n-space E^n can be strictly separated by a hyperplane if and only if for each set T consisting of $n+2$ or fewer points from $P \cup Q$ there exists a hyperplane which strictly separates $T \cap P$ from $T \cap Q$.

It is one of the inequities of mathematical history that Kirchberger is not given more credit for this basic theorem. Indeed, the fundamental theorem of Caratheodory for finite sets (which was not published until four years later [1]) is included as a special case where one of the sets consists of a single point. Even the remarkable theorem of Helly (about twenty years later [2]) follows readily from the work of this "forgotten" mathematician.

In his book on convexity, F. A. Valentine posed the problem of replacing the "separating hyperplanes" of Kirchberger's theorem by "separating spherical surfaces" [4]. E. G. Straus has shown that for $n = 2$, the critical number is five. The following theorem generalizes this result to Euclidean spaces of arbitrary finite dimension.

THEOREM. *Let P and Q be two finite subsets of E^n. Suppose for each subset T of $n+3$ or fewer points of $P\cup Q$ there exists a spherical surface which strictly separates $T\cap P$ from $T\cap Q$. Then there exists a spherical surface which strictly separates P from Q.*

Proof. Embed E^n in E^{n+1} and let Ω be an n-dimensional unit sphere in E^{n+1} which is tangent to E^n at an arbitrary point p. Let π be the stereographic projection of E^n onto Ω based at the point antipodal to p in Ω. Suppose that given any subset T of $n+3$ or fewer points of $P\cup Q$, there exists a spherical surface S in E^n which strictly separates $T\cap P$ from $T\cap Q$. That is, given any subset $\pi(T)$ of $n+3$ or fewer points of $\pi(P\cup Q)$, there exists a spherical surface $\pi(S)$ on Ω which strictly separates $\pi(T\cap P)$ from $\pi(T\cap Q)$ on Ω. Now $\pi(S)$ is the intersection of some n-dimensional flat H with Ω, and H strictly separates $\pi(T\cap P)$ from $\pi(T\cap Q)$ in the $(n+1)$-dimensional space E^{n+1}. Thus by Kirchberger's theorem, there exists an n-dimensional flat H_0 in E^{n+1} which strictly separates $\pi(P)$ from $\pi(Q)$. It follows that $\pi^{-1}(H_0\cap\Omega)$ is a spherical surface in E^n which strictly separates P from Q.

References

1. C. Caratheodory, Über den Variabilitätsbereich der Koeffizienten von Potenzreihen, die gegebene Werte nicht annehmen, Math. Ann., 64 (1907) 95–115.

2. E. Helly, Über Mengen konvexer Körper mit gemeinschaftlichen Punkten, Jber. Deutsch. Math.-Verein., 32 (1923) 175–176.

3. P. Kirchberger, Über Tchebychefsche Annäherungsmethoden, Math. Ann., 57 (1903) 509–540.

4. F. A. Valentine, Convex Sets, McGraw-Hill, New York, 1964, p. 177.

RESEARCH PROBLEMS

EDITED BY RICHARD GUY

In this Department the Monthly presents easily stated research problems dealing with notions ordinarily encountered in undergraduate mathematics. Each problem should be accompanied by relevant references (if any are known to the author) and by a brief description of known partial results. Manuscripts should be sent to Richard Guy, Department of Mathematics, Statistics, and Computing Science, The University of Calgary, Calgary 44, Alberta, Canada.

MONTHLY RESEARCH PROBLEMS, 1969–71

RICHARD GUY, University of Calgary, and
VICTOR KLEE, University of Washington

The posing of good unsolved problems is a difficult art. There are many simply stated problems which are, from the experience of experts, very unlikely to be solved within the next generation. But experts can be wrong, and we hope to see the Four Color Conjecture settled even if we do not live long enough to learn

the status of the Riemann and Goldbach hypotheses, "Fermat's Last Theorem," and odd perfect numbers. On the other hand, "unsolved" problems may turn out not to be unsolved at all, or to be much more tractable than the author, editor, and referee thought. Since, in any case, it was not our desire to produce a mausoleum of unsolved and apparently insoluble problems, we do not apologize that some of the Research Problems published in the first three years have already been disposed of.

The Research Problems department does not publish *solutions*. Any solution submitted to another department of the MONTHLY will be judged on its own merits, and may be found unsuitable for MONTHLY readers even though the *statement* of the problem was (we hope!) suitable. Thus it is desirable to avoid duplication of effort and assist further advances in partially solved problems by updating the information in articles such as this. Readers are invited to contribute at any time to future articles by sending further references, preprints and offprints of partial or complete solutions, and other comments on previously published Research Problems, to the Editor of the department.

References to Research Problems in the MONTHLY are given in the form: author [year, page]; they appear in the text in chronological order but are not listed in the bibliography at the end. Other references are given in three different forms, the first two of which are listed at the end: author (year) for papers that have appeared in print or will soon appear and whose year of publication is definitely known to us, author (tbp) for those that are apparently to be published (though exact references and even final acceptance may be lacking at press time), and author (wrc) ("written communication") for those that are unlikely to be published in present form and those for which we have no knowledge of publication plans. The authors of these papers may be contacted for more detailed information.

Klee [1969, 54] discussed the old problem of whether, in the euclidean plane, a convex body can have two equichordal points. His list of references omitted an earlier discussion of the problem by Hadwiger (1955). Petty and Crotty (1970) have described some noneuclidean geometries in which a convex body *can* have two equichordal points. Klee asked whether, in the euclidean plane, a convex body can have two equireciprocal points. That was a foolish question, for the foci of an ellipse are easily seen to be equireciprocal points. However, he meant to ask whether any plane convex body other than an ellipse can have two equireciprocal points, and that problem appears still to be open.

Saying he had heard the problem from someone else but was unable to trace the source, Klee [1969, 80] asked whether every polygonal region is illuminable from some point. He later traced the problem to E. Straus in the early 1950's. For regions with a smooth boundary, the illumination problem is solved negatively by means of the example below, which is the modification by Peter Ungar (wrc) of an example of Penrose and Penrose (1958). That is, the region shown in Figure 1 is not illuminable from any of its points. The same example was discovered independently by Peter Kornya (wrc), and other illuminating comments on the illumination problem have been made by Dean Hickerson (wrc),

Robert Israel (wrc), and Joseph Zaks (wrc). By extending the idea of the above construction it is possible to produce, for each natural number k, a plane region R_k with smooth boundary such that R_k is not illuminable from any set of k points.

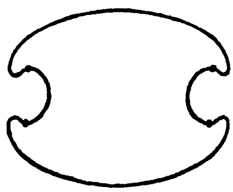


FIG. 1. (Upper and lower portions of the boundary are semiellipses with foci at the dark points. The exact shape of the central portion of the boundary is immaterial.)

Klee [1969, 286] asked what is the expected volume of a simplex whose vertices are chosen at random from a given convex body. Perhaps the most interesting unsettled case is that in which the body is a d-simplex for $d \geqq 3$. While the case $d=3$ could probably be handled by brute force, the case of a general simplex appears to be quite difficult. For spherical balls of arbitrary dimension the problem was solved by Kingman (1969), and more detailed information was obtained by Roger Miles (wrc).

Klee [1969, 288] discussed an old conjecture of R. D. Carmichael concerning the equation, $\phi(x)=n$. The table in Carmichael's 1908 paper was recently corrected by Wegner and Savitzky (1970), who tabulated all solutions x for all n up to 1978.

It was asked by Klee [1969, 408] whether the boundary of a d-dimensional convex body can contain segments in all directions. The answer was then known only for $d \leqq 3$. However, the question has now been answered negatively for $d \leqq 4$ by W. D. Pepe (wrc) and for all d by Ewald, Larman, and Rogers (1970). The condition of convexity cannot be abandoned entirely, for a construction of Bing (1961) can be modified so as to yield a 2-sphere in E^3 such that every line in E^3 is parallel to some segment in the 2-sphere.

Klee's question [1969, 539], "Is a body spherical if its HA-measurements are constant?", is of interest for several different notions of *body*. For some of them it has been answered negatively by Zaks (1971), but the case of convex bodies is still open. Firey (1970) has proved that, for all $1<m<d$, E^d contains nonspherical convex bodies of constant outer m-measure.

It was asked by Klee [1969, 678] whether all convex Borel sets can be generated in a Borelian manner within the realm of convexity. An affirmative answer was then known only for E^1 and E^2, and had also been obtained for cylindrical subsets of E^3 by Larman (1969). Since then, the problem and some of its relatives have been discussed by Rogers (1970) and a general affirmative solution for E^3 has been given by Larman (1971). However, Larman's methods do not appear to be adaptable to the 4-dimensional case, which remains open.

Another of Klee's problems [1969, 810] asked for an intrinsic characteriza-

tion of the intersection graphs of families of arcs in a circle. One was supplied by Tucker (1970), but there is still no good algorithm for testing whether a graph is of the sort in question. Also, no intrinsic characterization is known for the intersection graphs of families of plane convex sets. Renz (1970) is a reference given incompletely in [1969, 810].

Rosenthal [1969, 925] asked if almost commuting matrices are near commuting matrices. Luxemburg and Taylor (1970) answered the question affirmatively with the aid of nonstandard analysis, and more classical proofs were given by de Bruijn (1970) and W. Kahan (wrc).

Kennedy [1969, 1043] asked for an elementary proof of Peano's existence theorem for first-order differential equations. It turned out that one had been given by Grunsky (1960), and another was supplied by Walter (1971).

Kronk [1969, 1045] discussed the conjecture of Nash-Williams and Plummer that the square of every nonseparable graph is Hamiltonian. The conjecture has since been proved by Fleischner (tbpa), whose result was then sharpened by Hobbs (tbp). See also Fleischner (tbpb), (tbpc), (tbpd), and Fleischner and Kronk (tbp).

Duke [1969, 1128] asked whether the complete graph on $2n+1$ vertices can be packed with $2n+1$ copies of an arbitrary tree with n edges, and discussed the related problem (also due to G. Ringel) concerning the labeling of the vertices of such a tree with the integers $0, 1, \cdots, n$ so that the differences between adjacent labels take the values $1, 2, \cdots, n$ once each. Kotzig and Rosa (1970) have found some further infinite classes of trees for which the latter problem is solved, and the truth of the conjecture is established for all $n \leqq 14$. However, the vertex to be labeled 0 cannot be chosen arbitrarily. Two infinite classes of trees, one discovered independently by D. A. Sheppard, are given which contain a vertex that cannot be labeled 0. These contain all such trees with $n \leqq 9$, and the exceptional vertex in these cases is either terminal or the "center" of the tree.

Haggard and McWha (tbp) have obtained a sufficient condition for K_{2n+1} to be packed with a tree T in terms of T's adjacency matrix, but it lacks an algorithm to determine its applicability.

Branko Grünbaum (wrc) has suggested generalizing the problem to d-trees. A *d-tree* is a d-complex obtainable from the d-simplex by adding d-simplices, one at a time, just one new vertex being introduced at each stage. The conjecture is that for an arbitrary d-tree T with n d-faces and for $k \geqq 1$, the complete d-complex with $(d+1)kn+d$ vertices can be decomposed into

$$k\binom{(d+1)kn+d}{d}$$

copies of T. The labelling problem has also been generalized in an interesting article by Golomb (tbp). Klee [1970, 63] discussed the problem of determining $C(d, s)$, the maximum length of circuit codes of spread s in the d-dimensional cube. In Klee (1970) he gave a more detailed exposition of the same problem. A recent asymptotic result on $C(d, s)$ is due to Wyner (tbp).

Guggenheimer's problem [1970, 177] was concerned with the points inside a smooth convex curve C from which four normals to C can be drawn. Recent papers dealing with such points are those of Deo and Klamkin (1970) and of Heil (tbp), who makes a new conjecture. However, the original problem is still open.

The problem of Klee [1970, 288] concerning boundedness of isoperimetric ratios remains open. A reference given incompletely there is Larman and Mani (1970). Two other references which might well have been included are Kömhoff (1968) and (1970).

Ogilvy [1970, 388] raised some problems concerning iterated complex radicals. His conjectures have been recast by Shell (tbp) and some partial results concerning them obtained by Leon Gerber (wrc) and Hanns-Walter Rohde (wrc).

Subbarao [1970, 389] defined a *unitary perfect number* n to be equal to the sum of its *unitary divisors* (divisors d of n for which d and n/d are coprime, $d=1$ being included and $d=n$ excluded). Wall (tbp) discovered the unitary perfect number

$$2^{18}\cdot 3\cdot 5^4\cdot 7\cdot 11\cdot 13\cdot 19\cdot 37\cdot 79\cdot 109\cdot 157\cdot 313$$

and showed there are no smaller ones other than the four given earlier. Subbarao, Cook, Newberry, and Weber (tbp) have since shown that apart from those four every unitary perfect number is divisible by 2^{11} and by at least seven distinct odd primes. Subbarao (wrc) can also prove that the unitary perfect numbers have density zero.

The representation of Genocchi numbers conjectured by Gandhi [1970, 505] has been established by Carlitz (tbp), by Riordan and Stein (tbp), and by Edmund Pinney (wrc). Related results were obtained by Ralph James (wrc) and David Zeitlin (wrc).

Klee and Martin [1970, 616] asked whether a compact endset in E^3 must have measure zero. An affirmative answer in E^2 was established by Klee and Martin (1971) (and also, as it turned out, several years earlier by Roy O. Davies (wrc)), a negative answer in E^d for $d\geqq 4$ by Bruckner and Ceder (1971). Now the 3-dimensional case has been settled negatively by Larman (1971), who was also able to improve the Bruckner-Ceder construction in certain respects.

Now suppose that X is the boundary of a d-dimensional convex body and X_u is the union of the relative interiors of the maximal convex subsets of X. It was conjectured by Klee and Martin [1970, 616] that the $(d-1)$-measure of $X\sim X_u$ must be equal to zero. This was proved for $d\leqq 3$ in their 1971 paper, and has recently been proved for arbitrary d by Larman (tbp).

The problem of Freese, Miller, and Usiskin [1970, 867] to dissect a triangle into n triangles similar to it has been settled in the outstanding case $n=5$ by Usiskin and Wayment (tbp) and also by Ward Bouwsma (wrc), James Kiefer (wrc), Raymond Killgrove (wrc), Peter Mani (wrc), and Jim Morris (wrc). The dissection can be effected if and only if the triangle is right-angled or has angles of 30°, 30°, and 120°. The dissection of Figure 2 was discovered by several

authors in addition to those mentioned above. It is apparently unknown which triangles (if any) can be dissected into five triangles which are similar to each other but not to the large triangle.

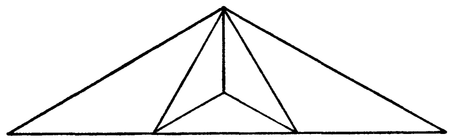


FIG. 2.

Fejes Tóth writes concerning his illumination problem [1970, 869] that a distribution of lamps constructed by A. Heppes makes it seem probable that Fejes Tóth's conjecture is false. Let m be positive and let the illumination function $f(x)$ be 1, $\frac{1}{2}(m+1)-x$ or 0 according as $0 \leqq x \leqq \frac{1}{2}(m-1)$, $\frac{1}{2}(m-1) < x \leqq \frac{1}{2}(m+1)$ or $\frac{1}{2}(m+1) < x$. Let P be a parallelogram with angle 45°, horizontal sides of length m, distance 1 apart, as in Figure 3. The illumination due to one lamp is given by the length of the intersection of P with a vertical line. Repeated translations of P through either of its two side vectors generate equidistant distributions of lamps of densities $1/m$ and 1, giving constant illumination 1 and m respectively. A combination of these gives a distribution of density $1+1/m$, and constant illumination $1+m$, which solves the lamp problem for $D=1+1/m$. If m is irrational the distribution is not periodic and cannot be the union of congruent lattices. To disprove the conjecture it remains to show that this solution is unique

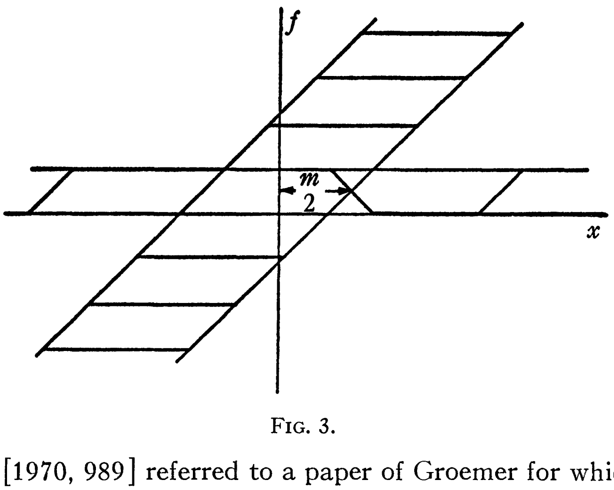


FIG. 3.

Chakerian [1970, 989] referred to a paper of Groemer for which incomplete details were given. The full reference is Groemer (1969).

The last reference of Wills [1971, 47] may now be given more precisely as Wills (1971). Edward A. Bender (wrc) writes that from a result of Steinhaus (1947) it follows that

$$s^*(m, 2) \leqq \frac{1}{2} + \frac{1}{2}\left(\frac{m-1}{\pi} + 1\right)^{1/2}$$

and from a result of Nosarzewska (1948) that

$$s(m,2) \leqq \frac{1}{4}+\frac{1}{2}\left(\frac{m-1}{\pi}+\frac{1}{4}\right)^{1/2},$$

where $s(m, 2)$ is the supremum considered by Wills and $s^*(m, 2)$ is the corresponding supremum for general (not necessarily convex) bodies.

M. Rosenfeld writes that an attack on his problem [1971, 49] of finding the number of graphs on n vertices with k cliques is being made for $k=3$ following a suggestion of P. McMullen. It is hoped to enumerate the Gale diagrams of the duals of the polytopes corresponding to the graphs by the Redfield-Pólya theorem.

R. S. Doran (tbp) has answered his question [1971, 178] affirmatively; R. F. Doser and E. A. Pedersen (wrc) have also proved that a *-algebra is symmetric if and only if the algebra obtained by adjoining an identity is symmetric.

The problem of Higgins and Ballew [1971, 274] on finite groups has received a number of solutions. Richard Brauer writes that the essential idea involved was quite familiar to Frobenius, certainly after Schur's thesis of 1901 and, he believes, even before. Other solvers include Bryant and Kovacs (1971), M. Benard (wrc), E. Formanek (wrc), J. S. Frame (wrc), R. Freese, C. Landauer, and J. McKay (wrc).

Hering's problem [1971, 275] on inequalities has been solved by Hering (tbp) himself.

In connection with his problem [1971, 385], Singmaster has extended his search to 2^{48} and shown (tbp) that the nontrivial occurrences of integers, appearing more than once as binomial coefficients, are still only 120, 210, 1540, 7140, 11628, 24310, and

$$3003=\binom{78}{2}=\binom{15}{5}=\binom{14}{6}.$$

He gives an infinity of solutions of

$$\binom{n+1}{k+1}=\binom{n}{k+2}$$

in the form $n=u_{2i+2}u_{2i+3}-1$, $k=u_{2i}u_{2i+3}-1$, where $u_0=0$, $u_1=1, \cdots$ are the Fibonacci numbers. Abbott, Erdös, and Hanson (tbp) write $N(t)$ for the number of times $t>1$ occurs as a binomial coefficient and prove that the average and normal order of $N(t)$ is 2, that $N(t)=O(\log t/\log\log t)$ and that $N(t)\geqq 6$ infinitely often. To obtain the second result they use Ingham's deep theorem on the existence of a prime between x and $x+x^{5/8}$, and note that if one assumes Cramer's conjecture that there is a prime between x and $x+(\log x)^2$, then their argument gives $N(t)=O((\log t)^{2/3+\epsilon})$.

Duke [1971, 386] conjectured that $\beta(G)\geqq 4\gamma(G)$ for each connected graph G, where β is the Betti number and γ is the genus. Nordhaus, Ringeisen, Stewart, and White (tbp) observe that the conjecture is correct when $\gamma(G)$ is 1 or 0, and

that for any G the conjecture would follow from the inequality $\gamma_m(G) \geqq 2\gamma(G)$, where $\gamma_m(G)$ is the *maximum genus* of G (the largest genus among surfaces in which G admits a 2-cell embedding). Duke later writes that Martin Milgram and Peter Ungar have independently disproved his conjecture.

R. L. Graham (wrc) sends a denser packing than that of Fejes Tóth [1971, 528] of parasites on the stem of a plant. The problem is equivalent to packing points in an infinite strip of unit width, no pair of points to be at distance less than 1. The Fejes Tóth packing (Figure 4) gave an average of one new point for each $(\sqrt{2}+1/\sqrt{3})/4 \approx 0.4979$ units of strip length. Graham's packing (Figure 5) gives one for each $(1+\sqrt{4\sqrt{3}-3})^{1/2}/6 \approx 0.4970$ units; he hopes to establish that this is optimal. He asks for the densest packing of points in a strip of width $w>0$. He observes that it follows from the work of N. Oler (see Folkman and Graham (1969)) that if w is a whole multiple of $\sqrt{3}/2$, then the obvious hexagonal packing is optimal; also that if $w<\sqrt{3}/2$ it is clear what to do. For other values of w, the problem is still open. There are analogous problems of *covering* (with unit circles). Graham's packing has been discovered independently by Gordon L. Miller (wrc).

Fink (tbp) has settled Herda's problem [1971, 888] by showing that a circle maximizes the minimum pseudo-diameter.

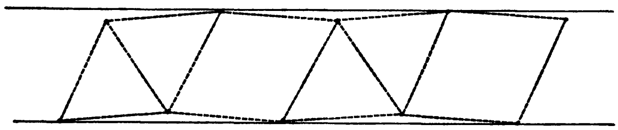


FIG. 4.

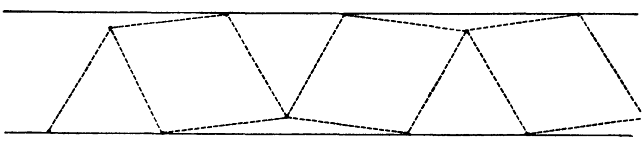


FIG. 5.

References

H. L. Abbott, P. Erdös and D. Hanson (tbp), *On the number of times an integer occurs as a binomial coefficient*, this MONTHLY.

R. H. Bing, *Tame Cantor sets in E^3*, Pacific J. Math., 11 (1961) 435–446.

A. M. Bruckner and J. Ceder, *A note on end-sets*, this MONTHLY, 78 (1971) 516–518.

N. G. de Bruijn, *Remark on a paper by W. A. J. Luxemburg and R. F. Taylor*, Notitie #44, (1970), 27:5:70, Dept. Math., Tech. Univ., Eindhoven.

R. M. Bryant and L. G. Kovács, *A note on generalized characters*, Bull. Austral. Math Soc., 5 (1971) 265–269.

L. Carlitz (tbp), *A conjecture concerning Genocchi numbers*, Norske Vid. Selsk. Skr. (Trondheim).

Narsingh Deo and M. S. Klamkin, *Existence of four concurrent normals to a smooth closed curve*, this MONTHLY, 77 (1970) 1083–1084.

R. S. Doran (tbp), *A generalization of a theorem of Civin and Yood on Banach *-algebras.*

G. Ewald, D. G. Larman and C. A. Rogers, *The directions of the line-segments and of the r-dimensional balls on the boundary of a convex body in Euclidean space*, Mathematika, 17 (1970) 1–20.

A. M. Fink (tbp), *A circle maximizes the minimum pseudo-diameter.*

W. J. Firey, *Convex bodies of constant outer p-measure*, Mathematika, 17 (1970) 21–27.

H. Fleischner (tbpa), *The square of every non-separable graph is hamiltonian.*

——— (tbpb), *The total graph of a block is hamiltonian.*

——— (tbpc), *On a certain spanning subgraph of a connected bridgeless graph.*

——— (tbpd), *On line-critical blocks.*

——— and H. V. Kronk (tbp), *Hamiltonsche Linien in Quadrat brückenloser Graphen mit Artikulationen*, Monatsh. Math.

J. Folkman and R. L. Graham, *A packing inequality for compact convex sets of the plane*, Canad. Math. Bull., 12 (1969) 745–752.

S. Golomb (tbp), *How to number a graph*, Graph Theory and Computing (ed. R.C. Read) Academic Press, New York, 1972.

H. Groemer, *On plane sections and projections of convex sets*, Canad. J. Math., 21 (1969) 1331–1337.

H. Hadwiger, *Ungelöste Probleme Nr. 3*, Elem. Math., (1955) 10–19.

G. Haggard and P. McWha (tbp), *Packing n-trees in* K_{2n+1}.

E. Heil (tbp), *Linienfamilien ebener konvexer Bereiche*, Tagung "Geometrische Ordnungen und Konvexer Körper," Oberwolfach, 29:6–3:7:70.

F. Hering (tbp), *Eine Verallgemeinerung der Ungleichungen vom arithmetischen und geometrischen Mittel*, Monatsh. Math.

A. M. Hobbs (tbp), *The square of a nonseparable graph is Hamiltonian connected*, Abstract, AMS Notices, 18 (1971) 553.

J. F. C. Kingman, *Random secants of a convex body*, J. Appl. Probability, 6 (1969) 660–672.

V. Klee, *The use of circuit codes in analog-to-digital conversion*, Graph Theory and its Applications (B. Harris, ed.) Academic Press, N. Y., 1970 pp. 121–131.

V. Klee and M. Martin, *Semicontinuity of the face-function of a convex set*, Comment. Math. Helv., 46 (1971) 1–12.

M. Kömhoff, *An isoperimetric inequality for convex polyhedra with triangular faces*, Canad. Math. Bull., 11 (1968) 723–727.

———, *On a 3-dimensional isoperimetric problem*, Canad. Math. Bull., 13 (1970) 447–449.

A. Kotzig and A. Rosa, *Magic valuations of finite graphs*, Canad. Math. Bull., 13 (1970) 451–461.

D. G. Larman, *On the convex generation of convex borel sets in* R^3, J. London Math. Soc., (2) 1 (1969) 101–108.

———, *The convex borel sets in* R^3 *are convexly generated*, J. London Math. Soc., (2) 4 (1971) 5–14.

——— (1971), *A compact set of disjoint line segments in* E^3 *whose end set has positive measure*, Mathematika, 18 (1971) 112–115.

——— (tbp), *On a conjecture of Klee and Martin for convex bodies*, Proc. London Math. Soc.

——— and P. Mani, *Gleichungen und Ungleichungen für die Gerüste von konvexen Polytopen und Zellenkomplexen*, Comment. Math. Helv., 45 (1970) 199–218.

W. A. J. Luxemburg and R. F. Taylor, *Almost commuting matrices are near commuting matrices*, Indag. Math., 32 (1970) 96–98.

E. A. Nordhaus, R. D. Ringeisen, B. M. Stewart and A. T. White (tbp), *A Kuratowski-type theorem for the maximum genus of a graph.*

M. Nosarzewska, *Évaluation de la différence entre l'aire d'une région plane convexe et le nombre des points aux coordonnées entières couverts par elle*, Colloq. Math., 1 (1948) 305–311.

L. Penrose and R. Penrose, *Puzzles for Christmas*, The New Scientist, (25 Dec. 1958) 1580–1581 and 1597.

C. Petty and J. Crotty, *Characterizations of spherical neighborhoods*, Canad. J. Math., 22 (1970) 431–435.

P. L. Renz, *Intersection representations of graphs by arcs*, Pacific J. Math., 34 (1970) 501–510.

J. Riordan and P. Stein, *Proof of a conjecture on Genocchi numbers*, Los Alamos preprint LA-DC 12764, (1971).

C. A. Rogers, *The convex generation of convex Borel sets in Euclidean space*, Pacific J. Math., 35 (1970) 773–782.

D. L. Shell (tbp), *More on the iterated radicals problem*, Math. Gaz.

D. Singmaster (tbp), *Repeated binomial coefficients and Fibonacci numbers*, Fibonacci Quart.

L. Solomon, *The solution of equations in groups*, Arch. Math. (Basel), 20 (1969) 241–247.

H. Steinhaus, *Sur un théorème de V. Jarnik*, Colloq. Math., 1 (1947) 1–5.

M. V. Subbarao, T. J. Cook, R. S. Newberry and J. M. Weber (tbp), *On unitary perfect numbers*, Delta.

A. Tucker, *Characterizing circular-arc graphs*, Bull. Amer. Math. Soc., 76 (1970) 1257–1260.

Z. Usiskin and S. G. Wayment (tbp), *Partitioning a triangle into 5 triangles similar to it*, Math. Mag.

C. R. Wall (tbp), *The fifth unitary perfect number.*

W. Walter, *There is an elementary proof of Peano's existence theorem*, this MONTHLY, 78 (1971) 170–173.

K. W. Wegner and S. R. Savitzky, *Solutions of $\phi(x) = n$, where ϕ is Euler's ϕ-function*, this MONTHLY, 77 (1970) 287.

J. M. Wills, *Gitterpunkte und Volumen-Oberfläche-Verhältnis konvexer Mengen*, Arch. Math. (Basel), 22 (1971).

A. D. Wyner (tbp), *Note on circuits and chains of spread k in the n-cube*, I.E.E.E. Trans. Computers.

J. Zaks, *Nonspherical bodies with constant HA-measurements exist*, this MONTHLY, 78 (1971) 513–516.

CLASSROOM NOTES

EDITED BY DAVID DRASIN

Manuscripts for this Department should be sent to Robert Gilmer, Department of Mathematics, Florida State University, Tallahassee, FL 32306. Notes are usually limited to three printed pages.

FLY AROUND A CIRCLE IN A WIND

L. B. WILLIAMS, Reed College

1. Introduction. One of the standard ground reference maneuvers for an aircraft is to fly a circular track in windy weather. A calm day is a rarity and, especially near an airport, accurate tracking is desirable, obviously. The VFR (visual flight rules) pilot learns to correct for wind drift by observing the terrain beneath him and then adjusting his rate of turn correspondingly. Such corrections are not small, and in a 90° turn, say, from base leg to final approach to the runway, the proper application of these corrections can make the difference between a smooth approach and a sloppy, or even unsafe one. We determine the magnitude of some of these wind drift corrections in this note.

2. The Air Path. The ground track is a circle, radius r miles, with center at point C. The aircraft holds on this track, at a fixed altitude in a steady east wind, w miles per hour, and turns to the left, see Figure 1.

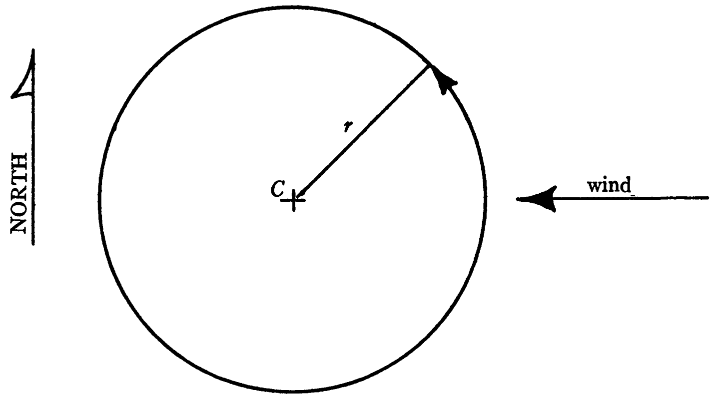


FIG. 1

Let this maneuver be observed from a free balloon, at the altitude of the aircraft, which is located above point C when the aircraft is at the east edge of the track. As time passes the aircraft will appear to an observer in the balloon to be flying a looped path (the "air path") and to be tracking a circle which is drifting off to the east, see Figure 2.

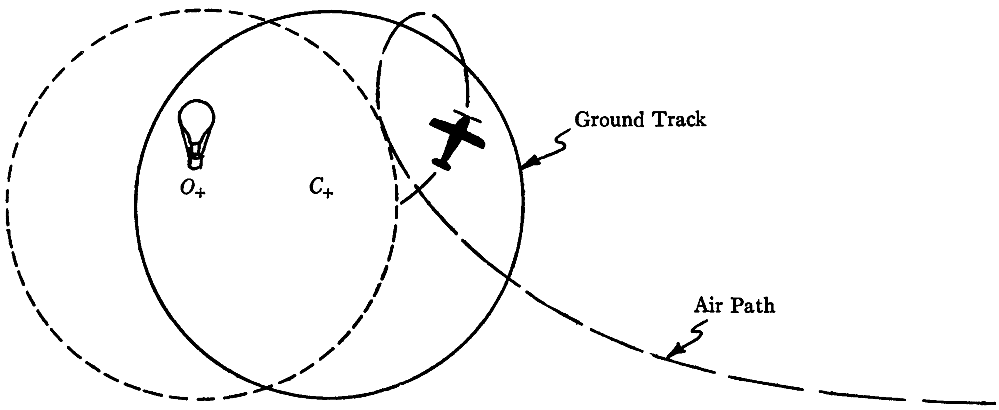


FIG. 2

We use a rectangular coordinate system whose origin O is at the balloon, the positive x-axis extending to the east, and the positive y-axis to the north. Let the aircraft fly at constant airspeed v miles per hour $(v>w)$, and let t be the time elapsed, in hours, since the balloon was above point C. Then the aircraft is at the position $P(x, y)$, and it has flown through an angle α (the "course angle") along the track. Parametric equations of the air path are

$$x = r\cos\alpha + wt, \qquad y = r\sin\alpha, \qquad t \geqq 0, \tag{1}$$

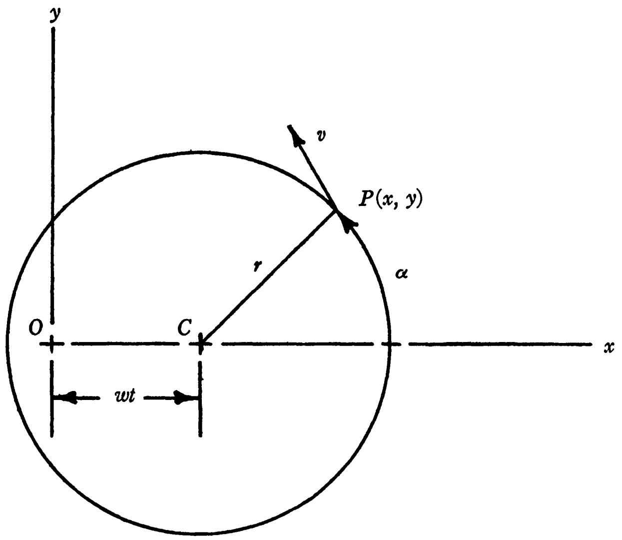


FIG. 3

in which x, y, and α stand for functions of t, see Figure 3.

Two parameters, α and t, appear in equations (1), and we derive a relation between them as follows (see equation (6) below). Let a time derivative be indicated by a prime. Then by differentiating equations (1) we see that

$$x' = -r\alpha' \sin\alpha + w, \qquad y' = r\alpha' \cos\alpha, \tag{2}$$

and the airspeed v is given by

$$v^2 = x'^2 + y'^2. \tag{3}$$

Elimination of x' and y' from (2) and (3) yields

$$r^2\alpha'^2 - 2wr\alpha' \sin\alpha + w^2 - v^2 = 0. \tag{4}$$

This equation in α' has two real roots, one positive and the other negative. Since $\alpha' > 0$, we have

$$r\alpha' = w\sin\alpha + \sqrt{(v^2 - w^2\cos^2\alpha)}. \tag{5}$$

Write $\alpha' = d\alpha/dt$, then the differential equation (5) can be expressed with variables separated in the form,

$$dt = \frac{r\,d\alpha}{w\sin\alpha + \sqrt{(v^2 - w^2\cos^2\alpha)}}\cdot$$

Rationalize the denominator on the right side, then substitute $s = \pi/2 - u$, and set $\alpha = 0$ at $t = 0$ to obtain

$$(v^2 - w^2)t = rw(\cos\alpha - 1) + rv\int_{\pi/2-\alpha}^{\pi/2} \sqrt{(1 - (w/v)^2 \sin^2 u)}\, du. \tag{6}$$

The definite integral in equation (6) is an elliptic integral of the second kind whose values are tabulated in available tables of integrals.

The time t to reach a course angle α, given by (6), can be used in equations (1) to locate the position, $P(x, y)$, of the aircraft on the airpath at that course angle. In noninstrument flying the course angle α is in fact the parameter a pilot uses to make his rate of turn corrections visually.

3. Time to reach a track position. The magnitude of the wind drift corrections is indicated by an example. Let $v = 100$, $r = 1$, and consider winds of 0, 20, 40, and 60 miles per hour. Times (in minutes) to reach quarter points on the track, assuming the airpath (1) is accurately flown, are approximated in Table I.

TABLE I. Minutes to reach course angle α

w \ α	90°	180°	270°	360°
0	.94	1.88	2.82	3.76
20	.87	1.74	2.85	3.96
40	.88	1.76	3.21	4.66
60	1.08	2.16	4.35	6.54

It takes about twice as long to fly around the circle in a gale as it does on a calm day.

4. Aircraft heading. A pilot usually checks his crab angle at the quarter points on the track. At $\alpha = 90°$ and $\alpha = 270°$, downwind and upwind respectively, the crab angle is zero. In a wind, the axis of the aircraft is not tangent to the ground track otherwise; however, it is tangent to the air path.

Let β be the heading of the aircraft at course angle α. The "heading" is the angle from north to the axis of the aircraft, clockwise looking down. Thus $\tan\beta = x'/y'$, and equations (2) give

$$\tan\beta = w/(r\alpha' \cos\alpha) - \tan\alpha.$$

Substituting for α' from (5), then rationalizing the denominator of the fraction, gives

$$\tan\beta = \frac{w\sqrt{(v^2 - w^2\cos^2\alpha)} - v^2 \sin\alpha}{(v^2 - w^2)\cos\alpha}. \tag{7}$$

Continuing the example of Section 3, we find that at the east edge of the track the headings (which can be read from the gyro-compass) for winds of 0, 20, 40, and 60 miles per hour are, respectively, 0°, 11°32′, 23°36′, and 36°52′.

The course angle at which the gyro-compass reads zero ($\beta=0$) is found from (7) to be given by $\cos \alpha = v/\sqrt{(v^2+w^2)}$. For the winds we used just above, a north heading is reached at the corresponding course angles 0°, 11°19′, 21°48′, and 30°58′, respectively.

5. Aileron control. In a coordinated turn, the setting of the ailerons is determined by the radius of curvature of the air path to be flown. Although this relationship is not linear, it is, at least, monotone.

The radius of curvature of a plane curve, given parametrically, is

$$R = (x'^2 + y'^2)^{3/2}/\left| x'y'' - y'x'' \right| .$$

We compute x'' and y'' from equations (2), and α'' by differentiating (4), using (5) to give α'. If the right member of (5) is represented by $f(\alpha)$, the radius of curvature R of the air path is given by

$$R = rv[1 - (w \sin \alpha)/f(\alpha)].$$

The radius of curvature of the airpath, at a given course angle, is proportional to the radius of the track. The amount of aileron control along a given track increases with increasing wind.

Using the same example as before, we compute approximate radii of curvature (in miles) at $\alpha=90°$ and $\alpha=270°$ for the four winds. These are shown in Table II.

TABLE II. Radius of curvature of the air path

w \ α	90°	270°
0	1	1
20	.67	1.56
40	.51	2.78
60	.36	6.25

For the most extreme case in this example (60 miles per hour wind) the aircraft is in a steep turn at the north point on the track, and practically is in straight and level flight upwind at the south point.

A CLASSIFICATION OF ULTRAFILTERS

R. J. ST. ANDRE, Central Michigan University

A standard exercise in introductory topology is to prove the existence of an ultrafilter that is not principal. The student considers $\mathfrak{F}_c$, the filter of all subsets

of an infinite set whose complements are finite, and shows any ultrafilter containing $\mathfrak{F}_c$ is not principal. It is interesting to note that this is the only type of ultrafilter which is not principal.

THEOREM. *Let $\mathfrak{F}$ be an ultrafilter on an infinite set X. Then either $\mathfrak{F}_c \subseteq \mathfrak{F}$ or $\mathfrak{F}$ is principal.*

Proof: Suppose $A \notin \mathfrak{F}$ for some $A \in \mathfrak{F}_c$. Then $X-A$ is a finite set and $X-A \in \mathfrak{F}$. There exists $x \in X-A$ such that $\{x\} \in \mathfrak{F}$ or else

$$A = \cap \{X - \{x\} : x \in X-A\} \in \mathfrak{F}$$

which is a contradiction.

Reference

1. Steven Gaal, Point Set Topology, Academic Press, New York, 1964, pages 258 and 266.

THE AXIOMS FOR EUCLIDEAN DOMAINS

KENNETH ROGERS, University of Hawaii

An integral domain D is usually defined to be **Euclidean** if:

(i) there is an integer-valued function g, defined on $D^* = D - \{0\}$, such that

(ii) for all a and b in D^*, we have $g(a) \leqq g(ab)$, and

(iii) for each a in D, and for each b in D^* which does not divide a, there exist q and r in D such that $a = qb + r$ and $g(r) < g(b)$. (The g-division algorithm.)

For instance, van der Waerden [**4**, p. 56], and Zariski and Samuel [**5**, p. 23] give axioms essentially equivalent to these. One may assume that $g \geqq 1$, since $g'(=g - g(1_D) + 1)$ has this property; indeed, Jacobson [**2**, p. 122] does this and also replaces (ii) by the stronger requirement that $g(ab) = g(a)g(b)$.

We show below that (i) and (ii) may be replaced by the simple condition:

(iv) the range of g is a set of integers bounded below.

This may be expected, since (iv) and the division algorithm are the only conditions used in the standard proof (as in Jacobson [**2**, pp. 121–123]) that Euclidean implies principal implies unique factorization. Thus, assumption (ii) is needed neither to insure that D is Euclidean nor in the modern proof on unique factorization. This latter point prompted Professor Basil Gordon to suggest to me that the status of (ii) should be clarified.

THEOREM. *If an integral domain D has a division algorithm with respect to an integer-valued function g, defined on D^* and bounded below, then D is Euclidean.*

Proof. The function g^* has properties (ii) and (iii), where

$$g^*(r) = \text{Min } \{g(rd) \mid d \in D^*\}.$$

For $g^*(a) \leqq g^*(ab)$, we need only note that

$$\{g(abd) \mid d \in D^*\} \subseteq \{g(ad) \mid d \in D^*\};$$

while for the g^*-division algorithm, taking a in D and b in D^*, with b not dividing a, we set $g^*(b) = g(bc)$ and apply the g-division algorithm to ac and bc. The remainder is a multiple of c, say $ac = q \cdot bc + rc$, with $g(rc) < g(bc)$. But $a = qb + r$, and

$$g^*(r) \leqq g(rc) < g(bc) = g^*(b),$$

as required.

We also wish to show that (ii) can be by-passed in deriving the existence part of the non-ideal-theoretic proof of the unique factorization theorem. Since condition (ii) is not needed for uniqueness, it could then be dropped, and the above theorem would just be an interesting exercise for the student.

Assume D satisfies the hypotheses of the theorem, but it contains **abnormal** elements, that is, non-units which cannot be expressed as a product of irreducible elements. Let a be an element of least g-value among all elements having an abnormal factor. Then a has an abnormal proper factor b, since abnormal elements are reducible with some factor abnormal, so we can apply the division algorithm and have $b = qa + r$, with $g(r) < g(a)$. Since b is an abnormal factor of r, the minimality property of a is contradicted. Note that the **existence** of irreducible elements was just proved also! Alternatively, any non-unit of least g-value is irreducible.

Anyone seriously wishing to study the question of different algorithm functions g for the same Euclidean domain should work through the definitive article on that subject, due to the late T. S. Motzkin [**3**]. In his consideration of **teas** (transfinite Euclidean algorithms), he also relaxes condition (ii). It seems appropriate to refer also to a paper of Hasse [**1**], where necessary and sufficient conditions are obtained for D to be a principal ideal domain: they amount to replacing (iii) by the condition that for all a in D and b in D^*, with b not dividing a, there exist m and n in D, such that

$$g(na - mb) < g(b).$$

The **sufficiency** is proved in the same way as in showing that Euclidean implies principal, so the interest lies in the conditions being **necessary** as well.

The author wishes to thank the UCLA Mathematics Department, where he spent a sabbatical leave during which this note was written, and, in particular, Basil Gordon for suggesting this work.

References

1. H. Hasse, Über eindeutige Zerlegung in Primelemente oder in Primhauptideale in Integritätsbereichen, J. Reine Angew. Math., 159 (1928) 3–12.
2. N. Jacobson, Lectures in Abstract Algebra I, Van Nostrand, Princeton, N. J., 1951.
3. T. S. Motzkin, The Euclidean Algorithm, Bull. Amer. Math. Soc., 55 (1949) 1142–1146.
4. B. L. van der Waerden, Modern Algebra I, Ungar, New York, 1949.
5. O. Zariski and P. Samuel, Commutative Algebra I, Van Nostrand, Princeton, N. J., 1958.

ON RIEMANN INTEGRABILITY

R. C. METZLER, University of New Mexico

The proof that a continuous function on a closed bounded interval is Riemann integrable is most often based on uniform continuity. Another common proof shows that the upper and lower integrals have equal derivatives in the intervals and then uses the mean value theorem for derivatives. It is interesting that a stronger theorem can be established which uses nothing more than elementary properties of upper and lower integrals, the order completeness of the reals, and the definition of limit. In addition this proof can be easily adapted to the Riemann-Stieltjes integral.

We assume that the upper and lower integrals,

$$\overline{\int_a^b} f(t)dt \quad \text{and} \quad \underline{\int_a^b} f(t)dt,$$

of a bounded function have been defined. We also assume the following elementary properties of upper and lower integrals:

(a) $$\overline{\int_a^b} f(t)dt \geqq \underline{\int_a^b} f(t)dt \qquad \textit{for each bounded function } f.$$

(b) $$\left.\begin{aligned} \overline{\int_a^b} f(t)dt &= \overline{\int_a^c} f(t)dt + \overline{\int_c^b} f(t)dt \\ \underline{\int_a^b} f(t)dt &= \underline{\int_a^c} f(t)dt + \underline{\int_c^b} f(t)dt \end{aligned}\right\} \quad \textit{for } a < c < b.$$

(c) *The upper and lower integrals are unchanged if f is replaced by a new function which differs from f at one point only.*

THEOREM. *Let f be a bounded function on the interval $[a, b]$. If f has a right-hand limit at each point of $[a, b)$, then f is Riemann-integrable on $[a, b]$.*

Proof: For any $\gamma > 0$, define

$$h_\gamma(x) = \begin{cases} \overline{\int_a^x} f(t)dt - \underline{\int_a^x} f(t)dt - \gamma(x-a) & a < x \leqq b \\ 0 & x = a. \end{cases}$$

It is enough to show that $h_\gamma(b) \leqq 0$ for all $\gamma > 0$, since this will imply

$$\overline{\int_a^b} f(t)dt \leqq \underline{\int_a^b} f(t)dt.$$

Suppose this is not true. Then there exists $\alpha > 0$ such that $h_\alpha(b) > 0$. Let $x = \inf \{t \in [a, b] \mid h_\alpha(t) > 0\}$. We shall show that the assumptions $h_\alpha(x) > 0$ and $h_\alpha(x) \leqq 0$ both lead to contradictions.

If $h_\alpha(x) > 0$, let $M = \sup \{|f(t)| : a \leqq t \leqq b\}$ and choose δ such that $0 < \delta < x - a$ and $2\delta M < h_\alpha(x)$. Then

$$\overline{\int_a^{x-\delta}} f(t)dt = \overline{\int_a^{x}} f(t)dt - \overline{\int_{x-\delta}^{x}} f(t)dt \geqq \overline{\int_a^{x}} f(t)dt - M\delta,$$

$$\underline{\int_a^{x-\delta}} f(t)dt = \underline{\int_a^{x}} f(t)dt - \underline{\int_{x-\delta}^{x}} f(t)dt \leqq \underline{\int_a^{x}} f(t)dt + M\delta,$$

so

$$h_\alpha(x - \delta) = \overline{\int_a^{x-\delta}} f(t)dt - \underline{\int_a^{x-\delta}} f(t)dt - \alpha(x - \delta - a) \geqq h_\alpha(x) - (2M - \alpha)\delta > 0.$$

This contradicts the fact that x is a lower bound.

If $h_\alpha(x) \leqq 0$, then, since $f(x+0)$ exists by assumption, we can choose δ such that $0 < \delta < b - x$ and $|f(x+0) - f(t)| < \frac{1}{2}\alpha$ for $0 < t - x < \delta$. By property (c) we can assume that $f(x) = f(x+0)$ for purposes of evaluating h_α. But then for $x < t < x + \delta$ we have

$$\overline{\int_a^{t}} f(y)dy = \overline{\int_a^{x}} f(y)dy + \overline{\int_x^{t}} f(y)dy \leqq \overline{\int_a^{x}} f(y)dy + [f(x + 0) + \tfrac{1}{2}\alpha](t - x),$$

$$\underline{\int_a^{t}} f(y)dy = \underline{\int_a^{x}} f(y)dy + \underline{\int_x^{t}} f(y)dy \geqq \underline{\int_a^{x}} f(y)dy + [f(x + 0) - \tfrac{1}{2}\alpha](t - x).$$

Therefore

$$\begin{aligned} h_\alpha(t) \leqq \overline{\int_a^{x}} f(y)dy + [f(x + 0) + \tfrac{1}{2}\alpha](t - x) - \underline{\int_a^{x}} f(y)dy \\ - [f(x + 0) - \tfrac{1}{2}\alpha](t - x) - \alpha(t - a) = h_\alpha(x) \leqq 0. \end{aligned}$$

This contradicts the fact that x is the greatest lower bound.

Clearly this theorem includes the theorems that continuous functions are integrable and that monotone functions are integrable. Given the "continuous almost everywhere" characterization of Riemann integrability, it shows that bounded functions with right-hand limits at every point of an interval are continuous almost everywhere on the interval. Of course the dual theorem on left-hand limits is an immediate consequence by consideration of $f(-x)$.

If g is a continuous nondecreasing function on $[a, b]$, we can use g as an

integrator function and define upper and lower Riemann-Stieltjes integrals. By defining

$$h_\gamma(x) = \overline{\int_a^x} f(t)dg(t) - \underline{\int_a^x} f(t)dg(t) - \gamma[g(x) - g(a)],$$

we can use the same proof to show that each bounded function with right-hand limits at every point of $[a, b)$ is Riemann-Stieltjes integrable with respect to g on the interval $[a, b]$. (The fact that g is assumed continuous allows us to change the value of f at one point without changing the upper or lower Riemann-Stieltjes integrals and also guarantees that the indefinite upper and lower integrals are continuous functions.)

MATHEMATICAL EDUCATION

EDITED BY J. G. HARVEY AND M. W. POWNALL

Material for this Department should be sent to either of the editors: J. G. Harvey, Department of Mathematics, University of Wisconsin, Madison, WI 53706; M. W. Pownall, Department of Mathematics, Colgate University, Hamilton, NY 13346.

A PLEA FOR A LIBERAL PROGRAM IN THE MATHEMATICAL SCIENCES

R. V. HOGG, University of Iowa

In the past it has been popular to classify persons, who have some interest in mathematics, as belonging to one of several categories: pure mathematician, applied mathematician, statistician, computer scientist, one interested in operations research (systems analysis or management science), etc. After acknowledging that there are persuasive reasons for the creation of specialties within the vast domain of the mathematical sciences, I would like to suggest that the common good will be advanced by a more *liberal* basic educational program in mathematics. Of course, the word *liberal* has several meanings: "generous," "charitable," "tolerant," "broad-minded," "progressive," "not bound by traditional ideas," and "in accord with the concepts of maximum individual freedom." But my thought is that we should train mathematicians to be liberal in *all* of these senses, for it is these characteristics that I would hope that many mathematicians would possess. Indeed I am convinced that possession of these characteristics by a significant number of the mathematicians of the future is a necessity if the subject is to remain vital and to interact profitably with other sciences.

Yet, as we know, there are some of us who do not qualify for such a label. But for the sake of our students, if not ourselves, it seems to me that we must

adopt a liberal attitude in constructing the future programs in the mathematical sciences. That is, these programs must be broad and progressive, and the teachers must be tolerant of other interests and in accord with the concepts of maximum individual freedom. And, frankly, I would like to see the idea of problem formulation and solution as the central theme of this program, in which the problems can range from the most abstract to those which deal directly with applications in the social, biological, or physical sciences. Today's students must have a choice, and they should be made aware of the various alternatives early in their college, or even high school, careers. Most of our present programs simply do not present the broad spectrum of possible outlets for the liberal mathematician. Hence many students, with strong interests in mathematics, will drift into other fields when, in truth, they are better suited for some work which comes under the large umbrella of the mathematical sciences.

As one who is interested in such a liberal program, I am distressed with the fact that we do not have a global mathematical organization covering the many facets of mathematics. So recently, I made a very limited historical study concerning the development of some of our mathematical organizations, and I would like to report my findings here and then conclude with a recommendation.

The American Mathematical Society grew out of the New York Mathematical Society, which was founded in 1888. One of the early presidents of AMS was Emory McClintock; and, in his 1894 presidential address "The Past and Future of the Society," which is found in the *Bulletin* [1], he states, "Finally, and I might say, above all, it is the object of the Society that every member should be stimulated to the most successful effort possible in his own branch of mathematical labor, whatever it may be; whether it be in teaching, or writing, or original investigation, or in any combination of these lines of activity. The investigator must also be a writer; the writer may present his own investigations, or comment upon or summarize or write the history of those of others or elaborate a treatise or textbook upon some special subject; but whoever may investigate, and whoever may write, it is the lot of almost all of us, in one way or another, to teach. For this reason it is plain that this Society is, and must always remain, a society of teachers. Any tendency to restrict its usefulness solely to the paths of investigation and publication should, for every reason of prudence and wisdom, be resisted. The management of any organization which does not commend itself to the great majority of those interested, must not indeed necessarily end in failure, but must certainly fail of producing the most appropriate, the most useful, and therefore the best results. While, however, expressing this general opinion, I would by no means be understood to disparage the work of the writers and investigators. Not every teacher, however successful, feels impelled to write for publication, and not every writer has time and facilities for original investigation; yet all of us take pride in such work when done by others, and all of us, as members of the Society, feel that it would fail of its highest objectives if it did not encourage in every way the production of good papers and books and, above all, the prosecution of original discovery."

These are words of a liberal mathematician. Please note his emphasis on the fact that the Society was one composed of teachers and "any tendency to restrict its usefulness solely to the paths of investigation and publication should . . . be resisted." It is interesting to note that THE AMERICAN MATHEMATICAL MONTHLY came into existence that same year, 1894. Of course, at that time, this journal was not sponsored by any organization; but its editors, in their opening statement in 1894 make clear the purpose of this publication. "It has seemed to the Editors that there is not only room but a real need for a mathematical Journal of the character and scope of this MONTHLY. At the present time there is no Mathematical Journal published in the United States sufficiently elementary to appeal to any but a very limited constituency, and that comes to its readers at regular intervals. Most of our existing Journals deal almost exclusively with subjects beyond the reach of the average student or teacher of Mathematics or at least with subjects with which they are not familiar, and little, if any space, is devoted to the solution of problems. While not neglecting the higher fields of mathematical investigation, THE AMERICAN MATHEMATICAL MONTHLY will also endeavor to reach the average mathematician by devoting regular departments to the important branches of Mathematical Science."

During the next 20 years, the MONTHLY began to stress collegiate mathematics more and more because there were other good organizations whose main interests were in elementary and secondary teaching. Then, in 1914, a request was made to the Council of the AMS to consider the feasibility of conducting, under the auspices of AMS, a journal for the field covered by the MONTHLY. In 1915, H. E. Slaught, in an article in the MONTHLY [2], explains what happens to that request. "Upon the presentation of this resolution at the meeting of the Council a committee of five was appointed to report to the Council. The report of this Committee was discussed at length by the Council at the usually large and representative meeting in New York, in April 1915, and finally the following resolution was passed with only two or three dissenting votes:

"It is deemed unwise for the American Mathematical Society to enter into the activities of the special field now covered by THE AMERICAN MATHEMATICAL MONTHLY; but the Council desires to express its realization of the importance of the work in this field and its value to mathematical science, and to say that should an organization be formed to deal specifically with this work, the Society would entertain toward such an organization only feelings of hearty good will and encouragement.

"While some members of the Council committee and some others in the Society feel that the Society might well broaden its scope of activity along the lines suggested and thus maintain its sphere of influence throughout the entire mathematical field, yet the decision of the Council was so emphatic as to leave no room for doubt concerning the present policy of the Society, both as to its own attitude toward the field of activity in question and as to the desirability of having this field provided for by an organization formed to deal specifically with this work."

The consequence of this action of the Council of AMS was the formation of the Mathematical Association of America in 1915 to support the field of collegiate mathematics, and the MONTHLY became the official journal of this new organization. I do not know whether or not the Council of AMS made the correct decision in 1915, but I do wonder what the situation would be today if there had been then greater support for the minority position "to broaden its scope of activity along the lines suggested and thus maintain its sphere of influence throughout the entire mathematical field."

There were, of course, other times at which the AMS could have reversed its position and taken a broader stance, but it selected not to do so. For illustration, in 1935, when the Institute of Mathematical Statistics was formed, a resolution, as reported on p. 227 in the *Annals of Mathematical Statistics*, 6 (1935), was passed "instructing the officers to investigate the feasibility of the affiliation of IMS with AMS or with the American Statistical Association." One of those officers, Allen T. Craig, in a personal communication, informed me that at that time it seemed that IMS was too mathematical for ASA and that AMS again did not want to spread out its wings to include this new organization. Of course, during the more recent years, we have seen other divisions of the mathematical community with the formation of organizations in applied mathematics, computer science, operations research, etc. Accordingly, in considering future programs in the mathematical sciences, we must recognize the situation as it is and make recommendations with this in mind.

In order to help construct the liberal program in mathematics to which I referred earlier, it seems that the existence of a global mathematical organization would be most important. One of its primary duties would be to open the channels of communication among the various divisions of our community. Obviously, I am not talking about communications related to details of research in the various specialties, but those presenting a bigger picture so that we can understand major developments and changes within the various groups. It is only with these interchanges that we can construct a liberal program that will reflect some of the excitement that is occurring in creative mathematics. These communications will help us better understand persons in different areas and, hopefully, this will promote tolerance which is necessary for this program. I must commend MAA for its efforts in this direction because, through this organization and its related activities, there has been an attempt to do much of this. I speak, however, of even a greater effort. Possibly this can be achieved through the strengthening of the Conference Board of the Mathematical Sciences. But, certainly the question of individual membership, at a fairly nominal cost, in some global mathematical organization must be explored. This could present the means of these communications. Such an organization could do much to encourage a rich, broad, and liberal program in the mathematical sciences at the elementary, secondary, collegiate, and graduate levels, a program that clearly presents to the students the various tracks, ranging from the abstract to the most applied mathematics. Obviously, there should be great flexibility and

all sorts of possible mixtures in this program, with the common goal of teaching problem solving at all levels. We must be tolerant, and recognize the fact that all mathematicians do not like the same thing (thank goodness). And yet we do have common elements and we do need a global organization to speak for our broad community. Accordingly, I urge that we join forces to make a massive effort for something that all of us can support: a better liberal mathematical education for the students of tomorrow.

References

1. E. McClintock, The past and future of the Society, Bull Amer. Math. Soc., 1 (1894–5) 85–94.

2. H. E. Slaught, The promotion of collegiate mathematics, this MONTHLY, 22 (1915) 251–53.

GROUPING BY ACADEMIC MAJOR IN COLLEGE CALCULUS

R. R. STEFFANI, Southern Oregon College

The Problem. Increasing enrollments coupled with a non-increasing number of staff members have made it imperative for many colleges to find ways of improving their mathematics program, particularly at the lower division level. Recent studies by Stocton [1] and Turner [2] have indicated that learning in a large lecture class is not inferior to that in smaller lecture classes. In many of the experiments with large classes, students were randomly assigned to smaller classes for problem solving, etc.

In view of these results a study was undertaken in 1969–1970 at Southern Oregon College to determine if a college calculus class could be structured to meet the needs of students pursuing various subject matter majors, yet at the same time not require additional staff members. A survey of the previous year's calculus classes indicated that approximately one-third were mathematics majors, one-third science and engineering majors, and one-third biology, business, and social science majors.

Experimental Procedure. The experimental calculus class was team taught two days per week in a large section of 75 students by three members of the mathematics department. On the remaining two days of the week, and alternating with the large sections, students grouped by academic major were taught in three small or "lab" sections. Forty-one students completed the three quarter sequence.

In the large section the basic concepts, theory, and skills of the calculus were presented, leaving little time for discussion or problem solving. In the "lab" sections concepts presented in the large section, applications of these concepts, new topics relevant to the subject majors at hand, and problem solving were considered.

The control group consisted of forty-seven students who completed the three term calculus sequence during the previous academic year in three traditionally taught, ungrouped classes.

Testing. Each quarter two teacher-made tests and one final were administered to the experimental group. In addition ten-minute quizzes were given during "lab" sessions. At the end of the three quarter sequence the *Cooperative Calculus Test* (CCT), Form B published by Educational Testing Service and an *Applied Problem Test* (APT) written by the author were given to the control and experimental groups.

Experimental Results. Analysis of covariance was used to test the null hypotheses of equal means for the control and experimental groups. The sum ($\bar{G}$) of the student's overall high school GPA, and the high school mathematics GPA was used as a covariate. The "F" test statistic in all cases indicated no significant differences at the five percent level. An analysis of covariance table for the control and experimental groups is included.

ANALYSIS OF COVARIANCE FOR CONTROL AND EXPERIMENTAL GROUPS

Criterion	Group Mean Score		df	F
	Control	Experimental		
N	47	41		
$\bar{G}$	5.68	6.28		
CCT	27.68	31.59	85	3.03*
APT	4.13	4.78	85	1.84*

* Not significant at five percent level.

In addition to comparing mean test scores of the control and experimental groups, null hypotheses of equal means were also tested for subgroups of both groups. No significant differences were found between subgroups of the control and experimental groups, nor were differences found between subgroups within the experimental group. However, the biology-business subgroup differed significantly from both the mathematics and science major subgroups within the control group in general calculus achievement as measured by the CCT. However, since the sample sizes were relatively small, replicative studies are being carried out to substantiate these differences.

Thus on the basis of the statistical analysis, it was concluded that relative to achievement in calculus, the control and experimental groups did not differ. However, subgroups within the control group did not perform at the same level, while subgroups within the experimental group attained the same level of achievement.

Regardless of the lack of significant differences indicated by the statistical analysis, our experience suggests that the use of academic major subgroups enabled the instructors to present unique material to his subgroup, and to delete or emphasize different aspects of other concepts. The students indicated that they liked to be exposed to the ideas and methods of different instructors. The two instructors not involved in teaching the large section also had more time to prepare material for their "lab" sections.

At the end of the academic year, the team members expressed their approval of the experimental procedure and agreed to continue it the following year, thereby providing an opportunity to carry out further and more detailed studies.

References

1. D. Stocton, An experiment with a large calculus class, this MONTHLY, 67 (1960) 1024–1025.
2. V. D. Turner *et al.*, A study of ways of handling large classes in colleges, this MONTHLY, 73 (1966) 768–770.

PROBLEMS AND SOLUTIONS

EDITED BY EMORY P. STARKE

All problems (both elementary and advanced) proposed for inclusion in this Department should be sent to E. P. Starke, 1000 Kensington Ave., Plainfield, N. J. 07060. Proposers of problems are urged to enclose any solutions or information that will assist the editors. Ordinarily, problems in well-known textbooks and results in generally accessible sources are not appropriate for this Department. No solutions (except those accompanying proposals) should be sent to Professor Starke.

ELEMENTARY PROBLEMS

Solutions of Elementary Problems should be sent to Problems Group, Mathematics Department, University of Maine, Orono, ME 04473. To facilitate their consideration, solutions of Elementary Problems in this issue should be typed (with double spacing) and should be mailed before March 31, 1972. Contributors (in the United States) who desire acknowledgement of receipt of their solutions are asked to enclose self-addressed stamped postcards.

E 2325. *Proposed by S. I. Rosencrans, Tulane University*

Prove that if $-1<\alpha<0$

$$\binom{2\alpha}{2n} \geqq (2n+1)\binom{\alpha}{n}^2, \qquad n = 0, 1, 2, \cdots,$$

while if $\alpha<-1$ the inequality is reversed.

E 2326. *Proposed by Harry Lass, California Institute of Technology*

Consider the following generalized ménage problem: N people labeled clockwise as $1, 2, 3, \cdots, N$, are seated at a table. If k people are chosen, labeled $1', 2', 3', \cdots, k'$, such that $1\leqq 1'<2'<3'< \cdots <k'\leqq N$, we desire at least α_1

individuals between $1'$ and $2'$, at least α_2 individuals between $2'$ and $3'$, etc., and at least α_k individuals between k' and $1'$.

Show that the number of such choices is

$$\frac{k\alpha_k + N - \alpha}{N - \alpha}\binom{N - \alpha}{k} \quad \text{with } \alpha = \sum_{i=1}^{k} \alpha_i.$$

E 2327. *Proposed by Kenneth Rosen, University of Michigan*

Let S_m^n be the sum of the reciprocals of the integers not exceeding m and relatively prime to n. Prove that for $m>n$, $n\geqq 2$, S_m^n is never an integer.

E 2328. *Proposed by D. R. Hayes, University of Massachusetts*

Suppose G is a semigroup having the property that, for every $a\in G$, there is a unique element $a^*\in G$ such that $aa^*a=a$. Prove that G is in fact a group.

E 2329. *Proposed by R. S. Luthar, University of Wisconsin*

Suppose that $0<a<1$ [so that $I=(0,a)$ is closed under multiplication]. (A) Find all continuous real-valued functions f defined on I which satisfy

$$f(xy) = xf(y) + yf(x).$$

(B) Find all continuous real-valued functions f defined on I which satisfy

$$f(xy) = xf(x) + yf(y).$$

E 2330. *Proposed by Richard Stanley, Massachusetts Institute of Technology*

Let f be a function from the positive integers to the integers satisfying $f(m+n)\equiv f(n) \pmod m$ for all m, $n\geqq 1$ (e.g., a polynomial with integer coefficients). Let $g(n)$ be the number of values (including repetitions) of $f(1)$, $f(2)$, $\cdots$, $f(n)$ divisible by n, and let $h(n)$ be the number of these values relatively prime to n. Show that g and h are multiplicative functions of n related by

$$h(n) = \sum_{d\mid n} \mu(d)g(d)(n/d) = n\prod_{p\mid n}\left(1 - \frac{g(p)}{p}\right).$$

SOLUTIONS OF ELEMENTARY PROBLEMS

Sequential Integers

E 2225 [1970, 308; 1971, 199]. *Proposed by Diane Comer and J. J. Tourneau, Fisk University*

Show that any positive integer S can be written in exactly k ways as the sum of two or more consecutive positive integers (in increasing order), where k is the number of positive odd divisors of S greater than 1.

II. *Comment by George Polya, Stanford University.* This problem is discussed in my *Mathematical Discovery*, vol. II (1965), p. 166, problems 15.48 and 15.49. The analogous 15.50 should also be noted.

A Permutation Automaton

E 2267 [1970, 1107]. *Proposed by Emilia Mycielska, University of California, Berkeley*

Given a permutation $a_0, a_1, \cdots, a_n$ of the sequence $0, 1, \cdots, n$; a transposition of a_i with a_j is called *legal* if $a_i=0$ for $i>0$, and $a_{i-1}+1=a_j$. The permutation $a_0, a_1, \cdots, a_n$ is called *regular* if after a number of legal transpositions it becomes $1, 2, \cdots, n, 0$. For which numbers n is the permutation $1, n, n-1, \cdots, 3, 2, 0$ regular?

Solution by M. T. Bird, San Jose State College. Let P denote the starting permutation, $P=[1, n, n-1, \cdots, 3, 2, 0]$, and let R denote the desired finishing permutation, $R=[1, 2, \cdots, n-1, n, 0]$. We observe that if $n=1$, then P is trivially regular. The proof is divided into three cases.

CASE 1: n is even. If n is even, say $n=2k$, then we observe that after $k-1$ legal transpositions, the permutation P becomes the following permutation:

$$[1, n, 0, n-2, n-1, n-4, n-3, \cdots, 4, 5, 2, 3].$$

No further legal transpositions are possible, and we conclude that P is not regular if n is even.

CASE 2: n is of the form $n=2^j-1$ for some $j\geqq 1$. For $i=1, 2, \cdots, j$ define the permutation P_i as follows:

$$P_i=[1, 2, \cdots, 2^i-1, 0; n-2^i+1, n-2^i+2, \cdots, n-1, n; n-2\cdot 2^i+1, n-2\cdot 2^i+2, \cdots, n-2^i-1, n-2^i; \cdots; 2^i, 2^i+1, \cdots, 2\cdot 2^i-2, 2\cdot 2^i-1].$$

The permutation P_i consists of $(n+1)/2^i$ "blocks" of 2^i members each; we indicate the separation of blocks by a semicolon. Note that the members within each block (except the first) are in increasing numerical order. We now note that P_1 is obtained from P by precisely $(n-1)/2$ legal transpositions, and that P_{i+1} is obtained from P_i by precisely $(n+1)/2$ legal transpositions for $i=1, 2, \cdots, j-1$. But $P_j=R$ and thus R is obtained from P by precisely $j(n+1)/2-1$ legal transpositions. Hence P is regular whenever n is of the form 2^j-1.

CASE 3: n is odd but not of the form 2^j-1. Then n can be written uniquely as $n=(2k+1)2^j-1$, where $j, k\geqq 1$. In this case we can define $P_1, P_2, \cdots, P_j$ as in Case 2; also as in Case 2, P_j is obtained from P by precisely $j(n+1)2-1$ legal transpositions. However, in this case, $P_j\neq R$; but after exactly k further legal transpositions on P_j, we obtain the following permutation:

$$[1, 2, \cdots, 2^j-1, 2^j; n-2^j+1, n-2^j+2, \cdots, n-1, n; 0, n-2\cdot 2^j+2, \cdots, n-2^j-1, n-2^j; \cdots; 2\cdot 2^j, 2\cdot 2^j+1, \cdots, 3\cdot 2^j-2, 3\cdot 2^j-1; 3\cdot 2^j, 2^j+1, \cdots, 2\cdot 2^j-2, 2\cdot 2^j-1].$$

No further legal transpositions are possible, and we conclude that P is not regular.

All numbers are now accounted for and we infer that P is regular if and only if $n=2^j-1$ for $j=1, 2, \cdots$.

Also solved by W. D. Bouwsma, Norman Miller, and J. A. Painter.

Completely Multiplicative Function

E 2268 [1970, 1107]. *Proposed by Leonard Carlitz, Duke University*

An arithmetic function $f(n)$ is *completely multiplicative* if $f(ab)=f(a)f(b)$ for all positive integers a, b. Show that

$$\sum_{ab=n} f(a)f(b) = \delta(n)f(n) \qquad (n = 1, 2, \cdots),$$

where $\delta(n)$ is the number of divisors of n, if and only if $f(n)$ is completely multiplicative.

I. *Solution by Harald Niederreiter, University of Southern Illinois.* If f is completely multiplicative, the result is immediately evident. Conversely, suppose that the condition is satisfied. Take $n=1$, and it follows that $f(1)=0$ or $f(1)=1$. Now suppose that $n \geqq 2$; let $n=p_1^{e_1} \cdots p_s^{e_s}$ be the canonical factorization of n and put $\alpha(n)=e_1+ \cdots +e_s$. It suffices to show that $f(n)=f(1)f(p_1)^{e_1} \cdots f(p_s)^{e_s}$ for all $n \geqq 2$. We proceed by complete induction on $\alpha(n)$. If $\alpha(n)=1$, then n is a prime (say $n=p$) and the proposition follows from the fact that

$$2f(p) = \delta(p)f(p) = f(1)f(p) + f(p)f(1) = 2f(1)f(p).$$

Suppose then that the proposition has been shown for all n with $\alpha(n) \leqq k$ where $k \geqq 1$. Take any n with $\alpha(n)=k+1$. Then

$$\delta(n)f(n) = 2f(1)f(n) + \sum f(a)f(b),$$

where the sum is taken over all a, b with $ab=n$ and $1<a$, $b<n$. It follows that $\alpha(a) \leqq k$, $\alpha(b) \leqq k$ so that the inductive assumption applies and thus

$$\delta(n)f(n) = 2f(1)f(n) + (\delta(n) - 2)f(1)^2 f(p_1)^{e_1} \cdots f(p_s)^{e_s}.$$

Since n is not prime, certainly $\delta(n)>2$ and so, for both $f(1)=0$ and $f(1)=1$ we get the desired result.

II. *Comment by C. S. Venkataraman, Sree Kerala Varma College, Trichur, S. India.* Attention is invited to Theorem 1 of J. Lambek, *Arithmetical Functions and Distributivity* (this MONTHLY, 73(1966) 969–973).

Also solved by Bennett College Problem Solving Team, S. J. Benkoski, D. M. Bloom, W. D. Bouwsma, B. R. Caine, Frederick Carty, M. K. Chowdury, Deborah Gale, Richard Gisselquist, M. G. Greening (Australia), Emil Grosswald, J. M. Jensik, Wells Johnson, S. A. Kalikow, David Kelly, Alfred Kohler, Joel Levy, D. J. Lohuis, Graham Lord, O. P. Lossers (Netherlands), D. E. Manes, Leon Mattics, A. J. Patsche, Bob Prielipp, Simeon Reich (Israel), Kenneth Rosen, T. Šalát (Czechoslovakia), E. F. Schmeichel, R. E. Spaulding, Jim Tattersall, Gregg Testini, A. M. Vaidya (India), J. A. Ventress, Mark Yu, K. L. Yocom, and the proposer. Partial Solutions by Anders Bager (Denmark), and Peter Kornya.

Rencontres mod 13

E 2269 [1970, 1107]. *Proposed by S. R. Conrad, Bayside, New York*

A shuffled deck of ordinary playing cards is dealt out in the manner of the French gambling game of "treize" [cf. Rouse Ball, *Mathematical Recreations and Essays*, p. 325]. The dealer deals out the first 13 cards as he calls the numbers $1, 2, \cdots, 13$ [J, Q, K counting as 11, 12, 13, respectively]. He repeats the procedure three more times, without replacement, exhausting the deck. What is the probability of (a) no "match"? (b) exactly k matches? (c) at least k matches? [A match is defined as calling out the number n while dealing card n.]

Editorial Comment. This problem is not new and has been considered several times in the literature. References were received from N. S. Mendelsohn (University of Manitoba), R. E. Greenwood (University of Texas), and L. G. Mossberg (Trollhättan, Sweden). Three incorrect solutions were received.

Mendelsohn refers to I. Kaplansky, *On a generalization of the "Problème des rencontres"* (this MONTHLY, 46(1939), 159–161). In this paper, Kaplansky derives a symbolic formula for the probability of no matches under quite general conditions. He gives no numerical results, however. Greenwood refers to Kaplansky, *Symbolic solution of certain problems in permutations* (Bull. Amer. Math. Soc., 50(1944) 906–914). Mendelsohn also refers to John Riordan, *Introduction to Combinatorial Analysis* (Wiley, New York, 1958). On p. 175, Riordan gives a symbolic solution to the problem in terms of rook polynomials, but notes that "Direct calculation $\cdots$ of these, though straightforward, is formidable and is usually replaced by approximations which will be given later."

These formidable calculations have been performed. Greenwood refers to his paper, *Probabilities of certain solitaire card games* (Journal Amer. Stat. Assoc., 48(1953) 88–93), and Mossberg to his paper, *Några rencontre-problem* (Elementa, 51(1968) 33–38, in Swedish). A typical numerical result is that the probability of no matches is 0.016233; this can be compared to $e^{-4}=0.018316$, which is (approximately) the probability of no matches in four independent trials with thirteen cards each.

Cosets of the Commutator Subgroup

E 2270 [1970, 1107]. *Proposed by Peter Yff, American University of Beirut, Lebanon*

Let a, b, c be elements of a group G of odd order. If $a^2b^2=c^2$, prove that ab and c are in the same coset of the commutator subgroup K.

Solution by Dennis Bertholf, Oklahoma State University. The solution is a corollary to the following more general result:

THEOREM. *Let a, b, c be elements of a group G of order m. If n and m are relatively prime and $a^nb^n=c^n$, then ab and c are in the same coset of the commutator subgroup K of G.*

Proof. For $x \in G$, let $[x]$ denote the coset of x modulo K. Since m and n are relatively prime, there are integers s and t such that $sm+tn=1$. Therefore

$$[c] = [c]^{sm+tn} = [c]^{sm}[c]^{tn} = [c]^{tn} = [c^n]^t = [a^nb^n]^t.$$

Since G/K is abelian, $[a^nb^n] = [ab]^n$ and hence

$$[a^nb^n]^t = [ab]^{nt} = [ab]^{nt+sm} = [ab].$$

Therefore ab and c are in the same coset of G/K.

Also solved by Anders Bager (Denmark), Bennett College Problem Solving Team, J. C. Binz (Switzerland), D. M. Bloom, Frederick Carty, J. P. Comiskey, G. F. Corliss, J. R. Courville, Harold Donnelly, G. W. Fehlhaber, M. G. Greening (Australia), D. C. Haines, H. E. Heatherly, G. A. Heuer, C. V. Holmes, John Hoslett, Wells Johnson, David Kelly, Peter Kornya, J. J. Leeson, G. B. Levy & Oubre, H. S. Lieberman, N. S. Mendelsohn, J. R. Olivier, Dan Richman, Sister Janet Schillinger, L. W. Shapiro, John Stout, Earl Taft, W. L. Werner, E. T. Wong, Mark Yu, Stojaković Zoran (Yugoslavia), J. R. Weaver & C. M. Bundrick, and the proposer.

A Determinant of Primitive Powers

E 2271 [1971, 77]. *Proposed by E. F. Schmeichel, the College of Wooster*

Let ω be a primitive nth root of unity. Let $A=(a_{ij})$, where $a_{ij}=\omega^{ij}$ for $1 \leqq i, j \leqq n-1$. If B_k denotes the matrix that results upon replacing the kth column of A by a column of 1's, prove that $\det B_k = -\det A$ for $k=1, 2, \cdots, n-1$.

Solution by C. V. Heuer and G. A. Heuer, Concordia College. Note that $\det A = \omega \cdot \omega^2 \cdots \omega^{n-1} \det A_1$ where A_1 is a Vandermonde matrix whose determinant is nonzero since $\omega, \omega^2, \cdots, \omega^{n-1}$ are distinct. Hence $\det A \neq 0$. Furthermore, since

$$(\omega^k - 1)(\omega^{(n-1)k} + \cdots + \omega^k + 1) = \omega^{nk} - 1 = 0,$$

it follows that $\omega^k + \omega^{2k} + \cdots + \omega^{(n-1)k} = -1$ for $k=1, \cdots, n-1$. Hence the system of equations

$$\sum_{j=1}^{n-1} \omega^{ij}x_j = 1, \qquad i = 1, \cdots, n-1$$

has the (unique) solution $X=(-1, -1, \cdots, -1)$. Since A is the coefficient matrix of this system if follows from Cramer's rule that $\det B_k = -\det A$ for $k=1, 2, \cdots, n-1$.

One should note that the requirement that ω be primitive is not necessary since if ω is not primitive then each of A and B_k has two identical rows so that both determinants are 0.

Also solved by Ari Amikam (Israel), Bennett College Problem Solving Team, D. M. Bloom, D. J. Bordelon, Robert Breusch, H. F. Bunch, B. Carlat, John Coolidge, R. A. Gibbs, Benjamin Greenberg, M. G. Greening (Australia), Harry Lass, O. P. Lossers (Netherlands), Carolyn Mac-

Donald, W. D. Markel, R. B. McNeill, Robert Patenaude, Simeon Reich (Israel), S. M. Rohde, Jonathan Ryshpan, L. W. Shapiro, Michael Shimshoni (Israel), J. S. Shipman, W. W. Tom, K. L. Yocom, David Zeitlin, and the proposer.

Uniquely Fibonacci

E 2272 [1971, 77]. *Proposed by D. C. Kay, University of Oklahoma*

It is well known that if $x=F_n$, $y=F_{n+1}$, $z=F_{n+2}$ are three consecutive members of the Fibonacci sequence then, because of the identities

$$x+y=z, \qquad xz=y^2\pm 1,$$

the $y\times y$ square may apparently be dissected to form an $x\times z$ rectangle with a net loss or net gain of one square unit (see Eves, *Survey of Geometry*, v. 1, p. 267). Prove that the Fibonacci sequence provides the only integral realization of this puzzle.

Solution by Simeon Reich, Israel Institute of Technology, Haifa. According to L. E. Dickson, *History of the Theory of Numbers*, (v.II, Chelsea, New York, 1952, p. 412), A. Lévy (Bull. Math. Élém., 15 (1909–1910), 113–115) showed that if (a, b) is the least positive solution $\neq(1, 0)$ of $x^2+xy-ky^2=1$, where k is a positive integer, every solution (u, v) is given by

$$u+v\omega=(a+b\omega)^n, \qquad \omega^2-\omega-k=0.$$

Noting the special case of $k=1$ we see that $a=b=1$ so that every solution (u, v) of $x^2+xy-y^2=1$ is given by $u+v\omega=(1+\omega)^n=\omega^{2n}$ for some integral $n\geqq 1$, where $\omega^2=\omega+1$. It is known that this implies $u=F_{2n-1}$ and $v=F_{2n}$. Since (u, v) satisfies $x^2+xy-y^2=1$ if and only if $(v, u+v)$ satisfies $x^2+xy-y^2=-1$, the full result follows.

Also solved by C. A. Bridger, Frederick Carty, M. G. Greening (Australia), Michael Goldberg, Douglas Lind, Robert Patenaude, Kenneth Rosen, Jonathan Ryshpan, J. J. Tattersall, W. G. Wild, C. C. Yalavigi (India), and the proposer.

Bridger notes that an early reference to $xz=y^2\pm 1$ appears to be in *An explanation of an obscure passage in Albrecht Girard's commentary on Simon Stevin's works*, by R. Simpson in the Philosophical Transactions of the Royal Society of London, 48(I) (1753), 368–377. Dickson observes that R. W. D. Christie (Math. Quest. Educ. Times, 73 (1900), p. 71) solved $x^2+xy-y^2=\pm 1$ by use of continued fractions. For further references see Dickson, *loc. cit.*

ADVANCED PROBLEMS

All solutions of Advanced Problems should be sent to J. Barlaz, Rutgers—The State University, New Brunswick, N.J. 08903. Solutions of Advanced Problems in this issue should be typed (with double spacing) on separate, signed sheets and should be mailed before March 31, 1972. Contributors (in the United States) who desire acknowledgment of receipt of their solutions are asked to enclose self-addressed stamped postcards.

An asterisk (*) means neither the proposer nor the editors supplied a solution.

5826. *Proposed by Richard Tapia, Rice University*

Let X be a Banach space with dual X^*. Consider $B=\{(x_\alpha, y_\alpha)\colon \alpha\in A\}$

$\subset X \times X^*$ such that

$$\langle x_\alpha, y_\beta \rangle = \delta_{\alpha\beta} \text{ (Kronecker delta).} \tag{1}$$

Arsove and Edwards call B a generalized basis if in addition to (1) we have

$$\text{The linear span of } \{x_\alpha : \alpha \in A\} \text{ is dense in } X. \tag{2}$$

Davis calls B a dual generalized basis if in addition to (1) we have

$$\text{The linear span of } \{y_\alpha : \alpha \in A\} \text{ is dense in } X^*. \tag{3}$$

Prove or disprove: these two notions of bases coincide in Hilbert space.

5827.* *Proposed by Ronald Hirshon, Polytechnic Institute of Brooklyn*

Let B be a finite p group, $p>2$. Let Z, the center of B, be cyclic with w a generator of Z. Does there exist an automorphism ϵ of B such that $w\epsilon = w^j$ for some j with $j \not\equiv 1 \bmod p$? If the answer is not always yes, will ϵ exist if we assume there is a maximal subgroup of B not containing Z?

5828. *Proposed by D. P. Giesy, Western Michigan University*

Let $x \in (0, 1)$. Is there an enumeration $\{q_n\}$ of the rationals in (0, 1) such that $\sum_{n=1}^{\infty} q_n/2^n = x$? (See Problem 5700 [1970, 1018–9], especially the Editorial Note.)

5829. *Proposed by D. P. Giesy, Western Michigan University*

Q is a countable subset of (0, 1). Find necessary and sufficient conditions on Q that it have the property: For every $x \in (0, 1)$ there exists an enumeration $q_1, q_2, \cdots$ of Q such that $\sum_{n=1}^{\infty} q_n/2^n = x$. (See Problem 5700 [1970, 1018–9], especially the Editorial Note, also the preceding problem.)

5830. *Proposed by Leonard Carlitz and R. A. Scoville, Duke University*

Let α be a positive irrational number and put

$$\phi(z) = \sum_{n=1}^{\infty} z^{[\alpha n]},$$

where $[\alpha n]$ denotes the greatest integer $\leqq \alpha n$. Show that $\phi(z)$ has the unit circle for a natural boundary.

5831. *Proposed by Albert Wilansky, Lehigh University*

Let C be a convex closed set in a normed space such that $C + D_1 \supset D_{1+\epsilon}$. Must C have non-empty interior? (Here $D_r = \{x : \|x\| \leqq r\}$, $\epsilon > 0$.)

SOLUTIONS OF ADVANCED PROBLEMS

Parametrization of $u^n + v^n = 1$

5761 [1970, 1015]. *Proposed by A. S. Adikesavan, Regional Engineering College, Tiruchirapalli, India*

Define $u(x)$ and $v(x)$ by the relation $\{u(x)\}^n + \{v(x)\}^n = 1$, where $n \geqq 2$ is an

integer, $u(0)=0$, $v(0)=1$, and $du(x)/dx=v(x)$. Determine $u(x+y)$ in terms of $u(x)$, $u(y)$, $v(x)$ and $v(y)$.

Solution by Emil Grosswald, Temple University. By assumption, $u'(=v)$ exists, so that $u^{n-1}u'+v^{n-1}v'=0$, and also $v'=-u^{n-1}v^{2-n}$ exists (except for possible isolated zeros of $v(x)$). Consequently, also $u''=v'$ exists and, by successive differentiations, it is easy to show that u and v are infinitely differentiable (with the possible exception of a countable set). It may be difficult to solve the problem if u and v are only required to satisfy this condition, i.e., to be functions of a real variable belonging to the class C^∞. Instead, we shall make the stronger assumption that u and v are actually holomorphic (i.e., complex analytic) functions in the complex plane. For such functions the following theorem of P. Montel holds (for a proof, see, e.g., A. I. Markushevich, *Entire Functions* (translation by Scripta Technica, Inc.), Elsevier, New York, 1966): *If two entire functions $u(z)$ and $v(z)$ satisfy a relation of the form*

$$\{u(z)\}^n + \{v(z)\}^n = 1$$

with integral $n \geqq 2$, then, for $n \geqq 3$, both functions reduce to constants, while for $n=2$ they are necessarily of the form $u(z)=\sin(h(z))$, $v(z)=\cos(h(z))$, with $h(z)$ an entire function.

In the present case, if $n \geqq 3$, so that u and v reduce to constants, it follows from $v=u'=0$ that $v(0)=1$ cannot hold and the problem, as here formulated, has no solution.

If $n=2$, every solution is of the form $u(z)=\sin(h(z))$, $v(z)=\cos(h(z))$, and from $v(z)=\cos(h(z))=u'(z)=\cos(h(z))h'(z)$ it follows that $h'(z)=1$, whence $h(z)=z+c$. The conditions $u(0)=\sin c=0$ and $v(0)=\cos c=1$ lead to $c=2k\pi$ with integral k, so that the only solutions are $u(z)=\sin(z+2k\pi)=\sin z$, $v(z)=\cos(z+2k\pi)=\cos z$ and, consequently, $u(x+y)=u(x)v(y)+u(y)v(x)$.

Also solved by B. E. Frejer who analyzes the growth possibilities of u and v, treated as real functions of a real variable, to show that no addition law of the form $u(x+y)=H\{u(x), u(y), v(x), v(y)\}$ can exist when n is odd.

S. J. Greenfield refers us to a note by Fred Gross (this MONTHLY, 73 (1966), p. 1093, ff.) which contains additional references and proofs of extensions of Montel's theorem cited above.

Extremal Problem in a Normed Linear Space

5762 [1970, 1015]. *Proposed by D. A. Zave, UNIVAC, Roseville, Minn.*

Let X be a real or complex normed linear space with norm $\|\cdot\|$. Let $a_i \neq 0$, $i=1, 2, \cdots, n$ be scalars and let $p \geqq 1$. If $b_1, \cdots, b_n$ and s are fixed elements of X, compute

$$\inf\left\{\sum_{i=1}^{n} \|a_i x_i + b_i\|^p : x_1 + \cdots + x_n = s\right\}.$$

Solution by A. A. Jagers, Enschede, Holland. Denote by α_p the stated infimum. Let $p>1$ and $p^{-1}+q^{-1}=1$. Then by Hölder's inequality

$$\alpha_p^{1/p}\left[\sum_{i=1}^{n}|a_i|^{-q}\right]^{1/q} \geqq \sum_{i=1}^{n}|a_i|^{-1}\|a_ix_i+b_i\| \geqq \left\|s+\sum_{i=1}^{n}a_i^{-1}b_i\right\|,$$

where in both relations the equality sign holds if for all i

$$x_i = -a_i^{-1}b_i + a_i^{-q}\left[\sum_{i=1}^{n}|a_i|^{-q}\right]^{-1}\left(s+\sum_{i=1}^{n}a_i^{-1}b_i\right).$$

Hence

$$\alpha_p = \left[\sum_{i=1}^{n}|a_i|^{p/(1-p)}\right]^{1-p}\left\|s+\sum_{i=1}^{n}a_i^{-1}b_i\right\|^{p}$$

if $p>1$. In a similar way one proves that

$$\alpha_1 = \min\{|a_i| : 1 \leqq i \leqq n\}\left\|s+\sum_{i=1}^{n}a_i^{-1}b_i\right\|.$$

Also solved by Joel Levy, O. P. Lossers (Netherlands), and the proposer.

A Hyperbolic Integral Limit

5764 [1970, 1015]. *Proposed by M. L. Glasser, Battelle Institute, Columbus, Ohio*

Evaluate the limit

$$L = \lim_{\alpha\to\pi}\sin\alpha\int_0^\infty \frac{\sinh(\gamma x)}{\sinh(\pi x)}\cdot\frac{dx}{\cosh x+\cos\alpha}, \quad |\operatorname{Re}\gamma| < \pi+1.$$

Solution by O. P. Lossers, Technological University, Eindhoven, Netherlands. Let δ be an arbitrary positive number. It is easily seen that we may write:

$$\begin{aligned}
&\lim_{\alpha\to\pi}\sin\alpha\int_0^\infty \frac{\sinh(\gamma x)}{\sinh(\pi x)}\,\frac{dx}{\cosh x+\cos\alpha}\\
&= \lim_{\alpha\to\pi}\sin\alpha\int_0^\delta \frac{\sinh(\gamma x)}{\sinh(\pi x)}\,\frac{dx}{\cosh x+\cos\alpha}\\
&= \lim_{\alpha\to\pi}\sin\alpha\int_0^\delta \left\{\frac{\gamma}{\pi}+O(x^2)\right\}\frac{dx}{\cosh x+\cos\alpha}\\
&= \lim_{\alpha\to\pi}\frac{\gamma}{\pi}\sin\alpha\int_0^\delta \frac{dx}{\cosh x+\cos\alpha}\\
&= \lim_{\alpha\to\pi}\frac{\gamma}{\pi}\sin\alpha\int_0^\infty \frac{dx}{\cosh x+\cos\alpha}.
\end{aligned}$$

(In the above, $\alpha\to\pi$ is taken as $\alpha\to\pi-$.) We evaluate the last integral:

$$\int_0^\infty \frac{dx}{\cosh x + \cos\alpha} = 2\int_0^\infty \frac{e^x dx}{e^{2x} + 2e^x \cos\alpha + 1}$$
$$= \frac{2}{\sin\alpha}\left[\arctan\left(\frac{y + \cos\alpha}{\sin\alpha}\right)\right]_{y=1}^{y=\infty}.$$

Hence we find that

$$L = \lim_{\alpha\uparrow\pi} \frac{2\gamma}{\pi}\left\{\frac{\pi}{2} - \arctan\left(\frac{1 + \cos\alpha}{\sin\alpha}\right)\right\} = \gamma.$$

Also solved by Emil Grosswald, D. A. Hejhal, Oscar Ocelot, P. H. Young, and the proposer.

Note. The proposer's discovery of the limit is derived from the Fourier cosine transform

$$\int_0^\infty \frac{\cos xy dx}{\cosh x + \cos z} = \pi \csc z \frac{\sinh(zy)}{\sinh(\pi y)},$$

using Parseval's theorem and the inverse transform.

Functions with Two Periods

5766 [1970, 1115]. *Proposed by Morris Newman, National Bureau of Standards*

Let α, β be complex numbers linearly independent over the reals. Let $g(z)$ be an entire function, and suppose that there are constants a, b such that for all z, $g(z+\alpha) = ag(z)$, $g(z+\beta) = bg(z)$. Prove that constants A, B exist such that $g(z) = A \exp(Bz)$.

Solution by Leonard Carlitz, Duke University. Essentially this is a known result. A meromorphic function $g(z)$ that satisfies

$$g(z + \alpha) = ag(z), \qquad g(z + \beta) = bg(z)$$

is called an elliptic function of the second kind. It is known (R. Fricke, *Elliptische Funktionen*, vol. 1, p. 220) that such a function is of the form

$$g(z) = Ae^{Bz}\frac{\sigma(z - u_1) \cdots \sigma(z - u_m)}{\sigma(z - v_1) \cdots \sigma(z - v_m)},$$

where $\sigma(z)$ is the Weierstrass sigma-function and $A, B, u_1, \cdots, u_m, v_1, \cdots, v_m$ are constants. Since $g(z)$ is assumed to be entire, it follows immediately that $g(z) = Ae^{Bz}$.

Also solved by B. Averbach, P. R. Chernoff, Irving Gerst, R. Goldstein (England), M. G. Greening (Australia), Emil Grosswald, D. A. Hejhal, Joseph Hesse, A. A. Jagers (Netherlands), Benjamin Lepson, O. P. Lossers (Netherlands)., M. B. Villarino, W. C. Waterhouse, and the proposer.

Editorial Note. For the original problem with $g(z)$ entire, most solvers show that $g'(z)/g(z)$ is elliptic and since $g(z)$ is entire, $g'(z)/g(z)$ can only have poles with positive residues at the zeros of $g(z)$. This being impossible by a theorem on elliptic functions, $g(z)$) can have no zeros, and the conclusion $g(z) = Ae^{Bz}$ follows. The proposer's solution avoids this particular theorem by observing that $g(z)g(-z)$ has two independent periods α, β, is entire, and therefore (Liouville's theorem) is nonzero constant (unless $g \equiv 0$). The same becomes true of the logarithmic derivative g'/g and the

result of the problem follows. R. Goldstein refers us to Copson, *Theory of Functions of a Complex Variable*, 1955, pp. 365–366 for the theorems on the Weierstrass elliptic functions given in the solution above.

Cauchy's Criterion

5767 [1970, 1115]. *Proposed by Howard Jacobowitz, New York University*

Let $g(t)$ be a real valued function on the open unit interval. The Cauchy criterion can be expressed in the following form:

$$\lim_{t\to 0} \sup_{0<t_1,t_2<t} \{g(t_1) - g(t_2)\} = 0$$

implies that $g(t)$ has a finite limit as $t\to 0$.

Prove or disprove the following modification:

$$\lim_{t\to 0} \sup_{t<t_1,t_2<2t} \{g(t_1) - g(t_2)\} = 0$$

implies that $g(t)$ has a finite limit as $t\to 0$.

Solution by S. A. Kalikow, Case Western Reserve University. The function $g(t) = (-\ln t)^{1/2}$ is a counterexample;

$$\sup_{t<t_1,t_2<2t} \{g(t_1) - g(t_2)\} = (-\ln t)^{1/2} - (-\ln 2t)^{1/2}$$

$$= \frac{\ln 2}{(-\ln t)^{1/2} + (-\ln 2t)^{1/2}}$$

which approaches zero as $t\to 0$.

Also solved by J. M. Ash, S. L. Campbell, R. A. Christiansen, L. E. Clarke (England), D. K. Cohoon, Crist Dixon, R. J. Driscoll, Hal Forsey, Robert Heller, D. A. Hejhal, Dennis Henkel, Ellen Hertz, G. A. Heuer, K. E. Hirst (England), A. A. Jagers (Netherlands), David Kelly, Peter Kornya, N. J. Kuenzi, Joel Levy, Douglas Lind, O. P. Lossers (Netherlands), D. B. MacMillan, E. A. Memmott, P. J. Owens (England), Nicholas Passell, Jürg Rätz (Switzerland), M. A. Roondog, Steven Russ, R. K. Tamaki, W. C. Waterhouse, K. L. Yocom, and the proposer.

Levy and Kelly note that a bounded counterexample is also possible; take sin $g(x)$ and note the inequality $|\sin a - \sin b| < |a-b|$. Memmott finds an analogous situation for infinite series: If $\{f_n\}$ is an arbitrary sequence of integers, then there is a real divergent series for which

$$\lim_{n\to\infty} \sum_{i=n}^{f_n} x_i = 0.$$

REVIEWS

EDITED BY J. ARTHUR SEEBACH, JR. AND LYNN A. STEEN

with the assistance of the mathematics departments of St. Olaf and Carleton Colleges

COLLABORATING EDITOR FOR FILMS: SEYMOUR SCHUSTER

Printed materials for review should be sent to: Book Review Editor, American Mathematical Monthly, St. Olaf College, Northfield, MN 55057. Films and correspondence relating to films should be sent to Seymour Schuster, Carleton College, Northfield, MN 55057.

All unsigned material is written by the editors. A boldface capital **C** *in the margin indicates that a review is based in part on classroom use. Professors willing to write such a review should inform the editor in order to avoid duplication.*

C *Topics in Ring Theory*. By Jacob Barshay. Benjamin, New York, 1969. 145 pp. \$17.50, \$7.95 (P). (Telegraphic Review, April 1970.)

The beginning chapters contain, with the exception of the theory of localization, more or less standard materials such as the isomorphism theorems, the Chinese remainder theorem, and the theory of euclidean-principal-ideal-unique factorization domains that are quite available in several existing standard undergraduate algebra texts. After a short chapter on exact sequences, the rest of the book is devoted to rings with chain conditions; this in the reviewer's opinion is the correct emphasis since most of the rings a student encounters in the undergraduate and graduate formal courses satisfy such conditions. All the results, to be sure, can be found in various textbooks; but they are usually embedded in the "bulky" books and hence not too appetizing for some students. The author has succeeded remarkably well in his choice of topics and theorems and organized them in a very readable and coherent form. The topics he has chosen are simultaneously useful in practice as well as fundamental to the theory that an aspiring young algebraist should master.

The style of writing is clear and the book is self-contained. Even though the chapters are short, plenty of material is included in some of them. For example, the chapter on noetherian rings contains the classical Hilbert basis, Krull intersection, Lasker-Noether decomposition, and Artin-Rees theorems. A separate chapter on Dedekind domains follows the material on noetherian rings. In a similar vein, after treating some of the general properties of Artinian rings, special mention is made of the semi-simple case leading to the Wedderburn structure theorem. There are sufficiently many exercises included at the end of each chapter.

This book is adaptable to a variety of purposes. It could be used as a text for one quarter or one semester course (as originally undertaken by the author) at the beginning graduate or advanced undergraduate level. It could also be used in a proseminar in conjunction with an algebra course, or it could be helpful in an individual reading project with faculty guidance where the number of credit hours is flexible.

J. S. HSIA, Ohio State University

C *Linear Algebra.* By Ross A. Beaumont. Harcourt, Brace, and World, New York, 1965. 216 pp. $5.50 (paper).

If an instructor wishes to pursue the study of linear algebra rigorously, including proofs of all the usual simple facts in the theories of vector spaces, matrices, systems of equations, and other standard topics, then Beaumont's *Linear Algebra* is a very fine textbook. The proofs and the motivation in general (including many examples and good exercises) are quite readable by good students. Moreover, the abstract vector space theory is well motivated in the first chapter, where the author develops the theory for the two and three dimensional spaces of geometrical vectors and shows them to be isomorphic to R^2 and R^3 respectively. Indeed, I would recommend the book for use as an introduction to abstract vector space theory, perhaps as part of a junior-senior year of abstract algebra, where the emphasis is to be placed on abstract vector spaces and one desires a sound exposition of the theory.

On the other hand, introductory courses in linear algebra are often expected to cover systems of linear equations, determinants, matrices, and other "useful" topics. If these topics are the main goal, an instructor may feel, as I do, that a rigorous exposition (including the time spent on abstract vector spaces) is at best unenlightening, no matter how well presented. Thus, except for a class presumed to have some feeling for linear algebra or to have the mathematical maturity to endure an abstract exposition, I would not recommend its adopting Beaumont's *Linear Algebra.*

D. R. SCRIBNER, University of Georgia

Topology for Analysis. By Albert Wilansky. Xerox College Publishing, Lexington, Massachusetts, 1970. xiii+383 pp. $13.50. (Telegraphic Review, June/July 1971.)

As its title implies, this book treats those aspects of topology which are of greatest use in analysis. The intended audience is at the advanced undergraduate or first year graduate level. Set theory is assumed known; for example, no definition of such terms as "countable" or "uncountable" is given.

The fourteen chapter titles will give some notion of the ground covered: Introduction; Topological Space; Convergence; Separation Axioms; Topological Concepts; Sup, Weak, Product, and Quotient Topologies; Compactness; Compactification; Complete Semimetric Space; Metrization; Uniformity; Topological Groups; Function Spaces; Miscellaneous Topics. The author uses "semimetric" in place of what other authors have called "pseudometric." In Chapter 3, *Convergence*, the author develops filters and nets simultaneously and throughout the book he continues to use these ideas side by side, showing their essential equivalence, but using whichever one is more natural in a given situation. The use of cardinal and ordinal numbers is virtually dispensed with; they do not occur until the last chapter. Instead the Stone-Čech compactification is used frequently as an example and counterexample.

The book has two interesting features. One is that minor remarks during the course of proof have their verifications enclosed in brackets, so that they

can be omitted on first reading. This device serves to clarify the main line of the argument and not to obscure it with too many details. The other feature is an extensive appendix which gives many of the standard counterexamples and theorems of topology in highly abbreviated form. This is done via 42 tables each headed by a property, e.g., table 5 is headed by "Completely Regular." There are two columns in each table, one headed "implied by," the other "not implied by." In table 5 one finds under the heading "Implied by" the notation "R. L θ5.3.5" indicating that a regular Lindelöf space is completely regular, the verification being via Theorem 5.3.5. Under "Not implied by" one finds "N 4.3 #3" indicating that a normal space need not be completely regular, the verification is via Exercise 3 of section 4.3. References are also made in these tables to standard examples and other authors. This latter feature should make the book quite useful as a handbook.

The format of the book is pleasant and quite readable. There appear to be very few typographical errors; the reviewer could find none, but knowing the vagaries of the printing process, he hesitates to aver that there are none. The index could be improved.

All in all this appears to be an excellent and useful addition to the literature of topology.

J. D. BAUM, Oberlin College

C In 1955, John Kelley asserted in the preface of his *General Topology* that he had with difficulty been prevented by his friends from labeling that book "What Every Young Analyst Should Know." Such has been the development of mathematical education that what was sixteen years ago partly a joke and partly a battle cry is now a truism, so that the book under review is quite properly titled "Topology for Analysis."

Those familiar with Professor Wilansky's first-rate *Functional Analysis* will approach this book with great expectations—and they will not be disappointed. Intended to serve the needs of both the beginning student and the mature mathematician, it fulfills both of these apparently disparate functions admirably. I list some of its excellences.

1. The level of exposition is consistently just right; always helpful, never either condescending or obscure.

2. The book conveys a feeling for the unity and continuity of mathematics, by its historical references and comments about the uses of topology in other areas.

3. Some 200 "examples" are scattered throughout the text. Some of these provide useful illustrations of the concepts; some apply theorems to special cases; some give instructive counterexamples to mark the limits of possible theorems; some refer to the literature for additional results.

4. The set theory bog, into which many texts settle at their outset, is avoided. The appropriate form of the maximal principle is stated when it is needed (to produce ultrafilters).

5. Metric spaces are handled in a useful and efficient way. Professor Wil-

ansky avoids both of the two opposite mistakes of (i) discussing them in detail before the general theory, necessitating doing many things twice; and (ii) discussing them at the end as an afterthought, depriving students of an important motivating example. They are introduced early as an example of topological spaces, and referred to continually as each new topological concept is introduced.

6. The practice of taking sides in the filters vs. nets skirmish (at the expense of the student, who needs to know about both to read the literature) is avoided. Filters, nets, and the connections between them are fully discussed, and use is made throughout the text of whichever is more useful in the particular circumstance.

7. There are some 1500 problems, graded by difficulty into three sets, most of them with hints and several with references to the literature. Among the more difficult problems appear such concepts and results as the Brouwer fixed point theorem, the Silverman-Toeplitz summability conditions, and paracompactness.

8. There are fifty pages of tables and notes, giving the text a useful handbook quality. Professor Wilansky, quoting Edwin Hewitt, disclaims interest in "whether every beta capsule of type delta is also a T-spot of the second kind." Still, the book implicitly recognizes that what is for one person a beta capsule may be for another a concept of crucial interest. One finds, for example, among the 39 items in the "σ-compact" table, each with a reference either to the text or the literature: "implied by LK.Σ.M." A dictionary provides the translations: a locally compact metrizable separable space is σ-compact.

I have only two demurrers. The book is a little too even, both in typography and (to a lesser extent) in exposition. The beginning student could well use more hints as to what is of primary and what of subsidiary importance. Also, the problems are served up in what seem to be overly small bites, at times with hints that are too helpful. Students may miss the useful experience of considering at greater length and working out for themselves a difficult multipartite problem. (This is, of course, a matter of individual taste.) As a final caveat, it should be noted that the book's title is to be taken seriously as defining its flavor and, to a large extent, its scope. Some topics of a more "geometric" nature are discussed briefly (local connectedness, arcwise connectedness, quotient topologies) or not at all (such results about the topology of Euclidean space as the Hahn-Mazurkiewicz theorem). In compensation, there are discussions in depth of such things as the rings of (real-valued) continuous functions on a topological space, the Stone-Čech compactification, the relation of sequential to topological properties, category, uniform spaces, and weak topologies. Two chapters provide brief but useful introductions to topological groups and function spaces, the latter coming to a climax with the Eberlein–Smulian theorem.

I made extensive use of page proofs of the text in a mixed undergraduate-graduate first course in topology in the fall of 1970. I covered about half of its material, moving quite quickly with a group of bright well-prepared students.

(The text became available at the end of the term, and students were urged to buy it.) In addition, I supervised a student who read about the same fraction of the page proofs as a reading course in the summer of 1970, and am presently (spring, 1971) repeating such a reading course. I have found the book eminently successful in both situations. I would strongly recommend it as a text for any one or two term first topology course (although it is perhaps more appropriate for a graduate than for an undergraduate course). Furthermore, it is a useful addition to any mathematician's library.

GARY LAISON, Lehigh University

Calculus with Analytic Geometry. By Burton Rodin. Prentice-Hall, Englewood Cliffs, N.J., 1970. 751 pp. $14.50. (Telegraphic Review, August–September 1970.)

This book, which received considerable advance billing as one of those written "for the student," is both innovative and paradoxical in its makeup. Innovative because it features fill-in "examples" as an educational device, gives non-"standard" proofs of many theorems, treats several classical topics in an unusual way, and presents material in a sequence that would surprise many; paradoxical because it is ultrarigorous in presentation yet not rigorous enough.

To see how the last statement is possible, imagine a book which goes out of its way to prove every little theorem on extrema and intermediate values, yet fails to explain the absence of the "$+C$" in integration by parts. It seems to this reviewer that the average interested reader at this level would be curious about the latter, but would rarely have the "maturity" to follow the former; yet this is typical of Rodin's book throughout, and one is left wondering, in a way, who the intended audience actually is: the ultrarigorous instructor, the ultra-passive student who avoids all theory (which generally appears in such small print as to almost discourage its reading) or some mystical in-between entity. Moreover, the book begins with finding areas under curves; the author explains that this is done to provide some natural motivation, yet this reviewer cannot help but feel that the plethora of sigma notation, diagrams (overwhelming in the section on volumes), and rigor (an understanding of induction is expected as early as p. 13) is enough to make more students want to drop the course (after all, they are taking three to five other courses) than it actually inspires.

The presence of non-"standard" proofs may be illustrated by two examples: the power rule for derivatives is proved by induction (the Binomial Theorem, strangely, is not even mentioned until near the end of the book) and the integral definition of natural logarithm is *derived* from the functional equation for the logarithm.

The psychological validity of the fill-in "examples" will be questioned by many (who may equate the fill-ins with a "mathematical coloring book"). For example, in

$$\frac{s(b) - s(a)}{b - a} = \underline{\qquad} = 30$$

(p. 145), how many would be tempted to write "30" in the blank space? And yet, if the book is to have reference value, consider the inconvenience of "skimming" through such extremes as the following:

$$\int_0^{3x} x^2 e^{xy} dy = \Big[\underline{\hspace{6em}}\Big]_{-=-}^{-=-} \quad \cdots$$

(p. 453). Given the bulk of "examples" and problems (including some very interesting ones), the multitude of theorems per page (from which it would seem difficult for the beginner to pick out the more "important" ones), and the text proper, there just seems to be too much to do to gain a *little* knowledge. In addition, the numbering (and lettering) system is somewhat cumbersome.

However, certain features show that the author really is interested in the student. Theorems are on occasion used first and proved second. Concepts are frequently illustrated in more than one notation, to show that these other notations do, in fact, exist. The fill-ins *seem* a step in the right direction. Where the fault may lie, then, may be in the author's major premise: that technique learning should be relegated to outside the classroom, since there is little time for it inside. Unfortunately, the "other things" which are to be done in class (such as some of the rigor in this book?) may be responsible for the very same student frustration the author is trying to eliminate.

It should be noted that arc length and surface area appear in a vector context. Vectors appear as n-tuples, rather than in i, j, k notation.

Much of the end of the book seems more suited for advanced calculus.

RICHARD REDFIELD, formerly Queens College (CUNY)

Ordinary Differential Equations. By Jack K. Hale. Wiley, New York, 1969. 348 pp. $14.95. (Telegraphic Review, April 1970.)

This excellent book should be in the library of every serious student of differential equations. The author has wisely stressed topics different from those in the well-known treatise by Coddington and Levinson, so the two texts are complementary rather than competitive, and they seem to be of comparable difficulty. Thus Green's functions, spectral theory, and boundary-value problems are not emphasized, while perturbation theory, periodic solutions, and Liapunov's method are treated quite fully. The book is enriched by a variety of nontrivial, interesting examples in the form of exercises.

Although the principal objectives are associated with applications, a high level of mathematical sophistication is maintained. For example, the Peano existence theorem is deduced from the Schauder fixed point theorem, an approach which smooths the way to modern theories of partial differential equations. A strong program in differential equations could be built around the use of this book together with a more standard text, followed, perhaps, by a treatise on control theory.

Since the author's style involves numerous cross references, it would have been better to provide section numbers and titles as running heads. Another

minor fault is that the author's desire for efficiency sometimes leads to a loss of historical perspective. For example, in the appendix on almost periodic functions, the name of Harald Bohr is omitted from the main text, and not even one of the various forms of the fundamental theorem is ever mentioned. Only a trifle more space would have sufficed to give a balanced view.

The bibliography is generally good, but the books of Pontryagin, Walter, and Siegel are better than some of the works that are cited, and should have been included. Siegel's book in particular should have been cited since, in the reviewer's opinion, Hale's book is a suitable introduction to it.

Page 50 indicates that the integral form of Bihari's inequality remains valid without the monotony condition. This is false, though the differentiated form does remain valid. A more serious slip occurs on p. 103, where it is suggested that the orbits in the case of a center are not ellipses. (The figure and accompanying text will doubtless be changed in a second printing.) A list of additional errors, mostly minor, can be obtained from the author.

To avoid misunderstanding let the reader return, now, to his pristine state of innocent expectation, and reread the first sentence of this review.

RAY REDHEFFER, University of California at Los Angeles

Lectures on Applications-Oriented Mathematics. By Bernard Friedman. Ed: Victor Twersky. Holden-Day, San Francisco, California, 1969. 268 pp. $15. (Telegraphic Review, February 1970.)

This book consists of revised and edited notes from a course of lectures given by Bernard Friedman during 1958–1965 at the Sylvania Laboratory in Mountain View. The book was prepared posthumously by friends and colleagues of Professor Friedman and is a symbol of their devotion to him. Indeed, they capture his style and flavor excellently as the change-of-pace quality of the book readily reveals.

The book is much more than a personal tribute; it is a precious morsel for anyone interested in the techniques and style of applied mathematics. Friedman was a rare kind of applied mathematician since he was also solid as a mathematician. This kind of applied mathematics is what is found in this book; a display of a set of techniques of applied mathematics. Always one has a confident feeling that the hands of a mathematician are at work. The level of sophistication varies, but material is always presented in a fluid style.

The book is somewhat more readable than Friedman's text "Principles and Techniques of Applied Mathematics" (Wiley, New York, 1956).

The topics treated are: Distributions, Spectral Theory of Operators, Asymptotic Methods, Difference Equations, Complex Integration, Symbolic Methods, Probability, and Perturbation Theory.

W. L. MIRANKER, I.B.M. Research Center, Yorktown Heights

TELEGRAPHIC REVIEWS

Telegraphic reviews are intended to give prompt notice of books, with sufficient information to assist in deciding whether to order an examination copy or suggest library purchase. Possible uses are indicated as follows:

B = college bookstore stock L = library purchase
P = professional reading S = supplementary reading
T = textbook E = teacher education
13 to 18 = freshman to second year graduate level usage
1 to 4 = approximate time in semesters to cover text
* = positive emphasis ? = negative emphasis

Books on high-school material (pre-calculus) are denoted REMEDIAL, and normally receive telegraphic reviews only if they are written for college students. Publishers are denoted by the standard abbreviations used in *Books in Print*, which gives complete addresses.

ALGEBRA, CATEGORY THEORY, ANALYSIS, P, L?, *K-Théorie*. Max Karoubi. Les Presses de l'Université De Montréal, 1971, 181 pp, $3.25 (P). A very quick "review" of classical K-theory is followed by a terse and categorical development of K-theory for Banach algebras. Most of the volume is devoted to categorical generalities extending classical K-theory. J.A.S.

ALGEBRA, LINEAR, T(14: 1), *Linear Algebra, Second Edition*. Serge Lang. A-W, 1971, xi + 400 pp, $10.50. A revised edition, with much of the material rewritten, new exercises and examples added and some new material introduced on unitary maps over the reals, the Jordan canonical form, and the spectral theorem. It is the author's intent that this text accompanied by his brief text on basic algebraic structures, would constitute a curriculum for an undergraduate algebra program. L.L.K.

*ALGEBRA, LINEAR ALGEBRA, T***(15-16), L, *Linear Algebra, Second Edition*. Kenneth Hoffman and Ray Kunze. P-H, 1971, viii + 407 pp, $12.50. A revision of the authors well-known *Linear Algebra*. "Our principal aim in revising...has been to increase the variety of courses which can easily be taught from it. ...we have structured the chapters...so that there are several natural stopping points along the way. ...we have increased the amount of material in the text." "...the basic philosophy behind the text is unchanged."R.J.

ANALYSIS, P, L, *Lecture Notes in Mathematics-170: Lectures in Modern Analysis and Applications III*. R.M. Dudley, J. Feldman, B. Kostant, R.P. Langlands, and E.M. Stein. Springer-Verlag, 1970, vi + 213 pp, $5.30 (P).Five lectures on modern harmonic analysis and applications taken from the seventh and eighth sessions of the lecture series *Modern Analysis and Applications* sponsored by the Consortium of Universities in Washington, D.C., and the University of Maryland during 1967-1969. R.J.

ANALYSIS, FUNCTIONAL ANALYSIS, P, L, *Numerical Ranges of Operators on Normed Spaces and of Elements of Normed Algebras: London Mathematical Society Lecture Note Series-II*. F.F. Bonsall and J. Duncan. Cambridge U Pr, 1971, 142 pp, $3.95 (P). The numerical range of a linear operator T on a normed space X, which reduces to

$\{ \langle Tx,x \rangle \mid \|x\| = 1 \}$ when X is a Hilbert space, was correctly defined only in 1961-62 independently by Lumer and Bauer. The notes give a well-documented exposition of the theory and include a survey of recent results and open problems. R.B.K.

CALCULUS, T(13: 2), S**, *Lectures on Freshman Calculus.* Allan B. Cruse and Millianne Granberg. A-W, 1971, xi + 641 pp, $10.50. These informal lectures do not attempt to treat difficult foundational questions (such as the definition of limit), but have presented calculus as a subject devoted to solutions of problems requiring derivatives and integrals. The authors draw a fine line between calculus and analysis with the observation that a mathematics major could get an analysis course the second year that would deal with the intricacies of continuity, compactness, convergence, and so on. The presentation of the topics as exciting problems to be solved does put the concepts into everyday settings. My reaction is that this is the way I like to present calculus while the student is working from a more standard text. L.L.K.

CALCULUS, T*(14: 2), *Calculus Two: Linear and Nonlinear Functions.* Francis J. Flanigan and Jerry L. Kazdan. P-H, 1971, xv + 443 pp, $10.95. A good blending of linear algebra and multivariable calculus. Implicit function theorems and change of variables in multiple integrals are missing and thereby weaken this as a second year text. Otherwise it is an interesting presentation with a good selection of problems. L.L.K.

DIFFERENTIAL AND INTEGRAL EQUATIONS, T(14-15), *Ordinary Differential Equations: Introductory and Intermediate Courses Using Matrix Methods.* H.K. Wilson. A-W, 1971, xiii + 377 pp, $10.95. The first six chapters cover introductory material in differential equations; the last four chapters (intermediate level material). Matrix methods are used. There is a table of course outlines using this book. Includes applications, many exercises, and hints for solutions and answers to problems. It gives a rather thorough development of introductory differential equations building upon prerequisite mathematics. R.J.

FOUNDATIONS, T(17: 1), S*, L*, *Introduction to Axiomatic Set Theory.* Jean-Louis Krivine. Humanities Pr, 1971, vii + 97 pp, $8.50. A concise, highly readable, graduate level text, revealing methods used in constructing relative consistency proofs, with Gödel's result for the continuum hypothesis being one of the examples. Especially helpful are the constant reminders, from the first page on, about the danger of confusing an intuitive concept (e.g. 'finite') with its formally defined counterpart, since in some models they will not correspond. Excellent introduction, a good translation from the French, but, alas, no index. P.F.

FOUNDATIONS, LOGIC, T(13: 1), *Introduction to Elementary Mathematical Logic.* Abram Aronovich Stolyar. Transl: Scripta Technica, Inc. Transl: and Ed: Elliott Mendelson. MIT Pr, 1970, vii + 209 pp, $5.95. Published in Russian in 1965, this text introduces propositional logic, propositional calculus and predicate logic to students in Soviet high schools (or other comparably prepared students). Exercises, historical references. L.A.S.

GENERAL, DICTIONARY, L. *Russian-English Dictionary of Electrotechnology and Allied Sciences.* Paul Macura. Wiley, 1971, x + 707 pp, $32.50. Almost entirely terms from technology (rather than science or mathematics) that gives quite thorough and careful coverage of its intended fields. It also includes many technical abbreviations and compounds, and some non-technical vocabulary of frequent appearance in technical writing. L.A.S.

NUMERICAL ANALYSIS AND APPLICATIONS, T?(16-17: 1). *General Dynamical Processes: A Mathematical Introduction. Mathematics in Science and Engineering, Volume 78.* Thomas C. Windeknecht. Acad Pr, 1971, xi + 179 pp, $9.50. A process is a set of mappings from a "time set" such as the nonnegative reals or positive integers into some set. It is a processor if the range of each mapping is a set of ordered pairs. A great deal of notation and terminology is introduced. The theorems are largely restatements of definitions. The lack of a symbol table and an inadequate index make the book difficult to read. R.B.K.

PHYSICS, S, P, L. *Lecture Notes in Physics-7. Lectures in Statistical Physics, From the Advanced School for Statistical Mechanics and Thermodynamics, Austin, Texas, U.S.A.* R. Balescu, J.K. Lebowitz, I. Prigogine, P. Résibois and Z.W. Salsburg. Springer-Verlag, 1971, v + 181 pp, $5.30 (P). Written notes of five lecture series, each series being given at the 1969 or 1970 school by one of the physicists named above. Prigogine, Salsburg, Résibois, and Lebowitz discuss entropy and dissipative structure, phase transitions, dynamical effects at the critical point in fluids and magnets, and exact results in equilibrium and non-equilibrium statistical mechanics, respectively, and Balescu gives an introduction to non-equilibrium statistical mechanics. D.F.A.

PROBABILITY AND STATISTICS, T(14: 1), S. *Elements of Decision Theory.* B.W. Lindgren. Macmillan, 1971, xii + 292 pp, $8.95. Introductory text requiring some mathematical maturity but no knowledge of calculus. An elementary background in probability theory would also be helpful. Concerned with basic principles rather than methodology. R.S.K.

PROBABILITY AND STATISTICS, P, L. *Lecture Notes in Mathematics-191: Séminaire de Probabilités V. Université de Strasbourg.* Ed: M. Karoubi and P.A. Meyer. Springer-Verlag, 1971, iv + 372 pp, $7.60 (P). Thirty-two papers from the 1969-70 Seminar, including ten by P.A. Meyer. Other contributors are Ph. Artzner, P. Assouad, J. Bretagnolle, R. Cairoli, P. Cartier, K.L. Chung, C. Dellacherie, C. Doleans-Dade, B. Maisonneuve, D. Revuz, J. de Sam Lazaro, J.B. Walsh, T. Watanabe, and M. Weil. All but four of the papers are in French. D.F.A.

PROBABILITY AND STATISTICS, T(16-17), S, L. *Nonparametric Statistical Inference.* Jean Dickinson Gibbons. McGraw, 1971, xiv + 306 pp, $11.95. This text treats most of the better known nonparametric techniques, assuming a background of a year course in probability and statistics. The presentation is theoretical, but not overly rigorous. Almost no numerical examples are included, and the problems are also theoretical. R.S.K.

PROBABILITY AND STATISTICS, ECONOMETRICS, P. *Lecture Notes in Operations Research and Mathematical Systems: Economics, Computer Sciences, Information and Control, Volume 40. The Coordinate-Free Approach to Gauss-Markov Estimation.* Hilmar Drygas. Springer-Verlag, 1970, viii + 113 pp, $3.50 (P). The author begins with a justification of the coordinate-free approach indicating its applicability in econometrics. He then develops the theory, applies it to linear unbiased estimators, and gives examples of its use. R.S.K.

REAL ANALYSIS, CALCULUS, T?(13-14: 2). *Elements of Calculus for Technical Students.* Lee W. Davis. Canfield Pr, 1971, viii + 312 pp, $10. The aim of this book is "to present elementary calculus as a tool to the technical student in today's expanding junior colleges and technical institutes." It contains a plethora of examples, problems, and applications and, as the author says, "theory and proofs are kept to a minimum or presented intuitively." They certainly are, and what's done isn't acceptable. Examples: the derivative of the sine function is determined by differentiating termwise its power series representation, which is called a "polynomial" since the author hasn't yet introduced infinite series; the proof (in the appendix) of the chain rule for differentiation is the one which can be supplied after a moment's reflection by any average calculus student; the Mean Value Theorem is never mentioned, yet graph sketching is discussed and all anti-derivatives of many functions are claimed to have been found; what's usually called "The Fundamental Theorem of Calculus" is taken here to be the definition of the definite integral of a function on an interval; in this text, to inquire about convergence of a sequence, one examines what happens to the nth term "when n approaches an infinitely large number." D.F.A.

*REAL ANALYSIS, FUNCTIONAL ANALYSIS, T(18), P, L**. *Bases in Banach Spaces I.* Ivan Singer. Grund. de Math. Wissenschaften, Band 154. Springer-Verlag, 1970, viii + 668 pp, $32.30. An outstanding contribution to the study of the basis problem in Banach spaces. The subject is brought up to date and unsolved problems are presented. This book should be in every mathematics research library. R.J.

*TOPOLOGICAL GROUPS, P, L. *Lie Algebras and Locally Compact Groups. Chicago Lectures in Mathematics.* Irving Kaplansky. U of Chicago Pr, 1971, xi + 148 pp, $2.50 (P). A "fifth account of Hilbert's fifth problem." Nice exposition with scattered exercises, comments, and bibliography giving some perspective on the problem and its solution. J.A.S.

Reviewers Whose Initials Appear Above

David F. Appleyard, Carleton; Paul Fjelstad, St. Olaf; Richard Jarvinen, Carleton; Lorraine L. Keller, St. Olaf; Roger B. Kirchner, Carleton; Richard S. Kleber, St. Olaf; J. Arthur Seebach, Jr., St. Olaf; Linda A. Seebach, St. Olaf.

Acknowledgments. The following have generously helped in evaluating books: Edward G. Begle, John Brothers, Harold Brown, Robert G. Buschman, Albert B. Farnell, Stephen Hechler, Henry Hiz, J. S. Hsia, M. A. Jenkins, Joseph Landin, Lucien LeCam, A. J. Lohwater, H. M. MacNeille, John McCarthy, Terrence W. Pratt, W. Rudin, Irving E. Segal, John Todd.

NEWS AND NOTICES

Edited by Raoul Hailpern, SUNY at Buffalo

Readers are invited to contribute to the general interest of this Department by sending news items to Mathematical Association of America, 1225 Connecticut Avenue, N.W., Washington, D. C. 20036. Items must be submitted at least two months before publication can take place.

PERSONAL ITEMS

Allegheny College: Dr. B. D. Haytock, Duquesne University, has been appointed Assistant Professor; Assistant Professor R. F. McDermot has been promoted to Associate Professor; Associate Professor C. A. Cable has been appointed Chairman of the Department of Mathematics.

Cornell University: Dr. D. R. Fulkerson of the Rand Corporation has been appointed the Maxwell M. Upson Professor of Engineering and Professor of Operations Research and Applied Mathematics; Professor W. F. Lucas has been appointed Director of the Center for Applied Mathematics.

Dr. O. O. Beck, Jr., Auburn University, has been appointed Assistant Professor at Florence State University.

Assistant Professor H. B. Keynes, University of Minnesota, has been promoted to Associate Professor.

Dr. J. H. Manheim, Bradley University, has been appointed Dean of the School of Letters and Science at California State College, Long Beach.

Professor Harry Siller, Hofstra University, died on April 15, 1971 at the age of 60. He was a member of the Association for thirty-two years.

DIRECTORY OF HISTORIANS OF MATHEMATICS

A world directory of historians of mathematics is being prepared by the Commission on History of Mathematics of the International Union for the History and Philosophy of Science. Scholars who are teaching or doing research in history of mathematics should communicate with the chairman of the Commission, Professor K. O. May, Department of Mathematics, University of Toronto, Toronto 181, Canada. Please send name and address (as you wish it for mail), a statement of your special fields of interest, and the languages you read.

The Commission expects to begin publication of an international journal of the history of mathematics in 1973. Meanwhile a newsletter will be distributed.

MATHEMATICAL ASSOCIATION OF AMERICA

Official Reports and Communications

APRIL MEETING OF THE IOWA SECTION

The fifty-eighth regular meeting of the Iowa Section of the MAA was held at Loras College, Dubuque, Iowa on April 23, 1971, Chairman Timothy Robertson presiding. Total attendance was fifty-four including forty-one members of the Association.

The meeting opened with the report that R. V. Hogg, State University of Iowa, was elected as Governor for the Iowa Section, and that Joseph Zimmerman, Hoover High, Des Moines, was selected for the annual Outstanding Teacher Award. The following officers were elected: Chairman, George Peglar, Iowa State University, Ames, Iowa; Vice-Chairman, Joseph Hoffert, Drake University, Des Moines, Iowa; Secretary-Treasurer, Basil Gillam, Drake University, Des Moines, Iowa.

The program consisted of two sessions, one in the morning followed the business meeting and adjourned at 10:50 A.M., and the afternoon session was held from 1:30–4:00 P.M. The morning session consisted of one paper:

Introducing mathematicians to some problems in robust estimation, by R. V. Hogg, University of Iowa, Iowa City.

The following program was presented in the afternoon:

Shapes of the future, by Victor Klee, University of Washington, Seattle (invited address).

Panel discussion on accreditation and certification. Panelists: William Waltmann, Wartburg College; Merlin Fisher, Iowa Central Community College; Marian Cornwall, Marshalltown Community College; E. W. Hamilton, University of Northern Iowa.

A questionnaire had previously been circulated, with very few returns. The members present were asked to complete the questionnaire after hearing the discussion. A résumé of the completed questionnaires indicated that the Iowa section of the MAA is very much opposed to any activity of the Association concerning these questions.

B. E. GILLAM, *Secretary-Treasurer*

APRIL MEETING OF THE TEXAS SECTION

The annual meeting of the Texas Section of the MAA was held at Midwestern University, Wichita Falls, Texas, on April 16 and 17, 1971. There were 184 people in attendance.

At the business meeting on April 17, the following officers were elected for the coming year: Chairman, Professor Louie Huffman, Midwestern University; Vice-Chairman, Professor Robert Northcutt, Southwest Texas State University; Regional Chairman of the MAA High School Contest, Professor W. S. McCulley, Texas A and M University; Secretary-Treasurer, Professor J. C. Bradford, Abilene Christian College.

Invited speakers were Professor Ward Cheney of the University of Texas who spoke on "A Survey of Approximation Theory" and Dr. A. B. Willcox, Executive Director of the MAA, who spoke on "England was Lost on the Playing Fields of Eton: A Parable for Mathematics."

Saturday morning a panel on Accreditation and Certification was held, followed by meetings of special interest groups. The following papers were presented in sessions meeting Friday afternoon:

1. *Arguments of proper values of normal transformations*, by A. Donnell and A. R. Amir-Moez, Texas Tech University.

2. *Hermitian transformations restricted to subspaces and inequalities among their eigenvalues*, by C. R. Perry, Jr. and A. R. Amir-Moez, Texas Tech University.

3. *L'Hospital's rule for matrices*, by Jean Richmond, Southern Methodist University.

4. *Density bounds for the sum of divisors function*, by C. R. Wall, P. L. Crews, and D. B. Johnson, East Texas State University.

5. *A generalization of a problem in number theory*, by M. G. Monzingo, Southern Methodist University.

6. *Some theorems on the semidirect product of groups*, by D. E. Edmondson, University of Texas at Austin.

7. *Addition chains*, by Robert Giese, University of Houston.

8. *Boolean near rings*, by J. R. Courville, University of Southwestern Louisiana.

9. *Research in near rings using a digital computer*, by Russell Schexnayder, University of Southwestern Louisiana.

10. *Commutator near rings*, by H. E. Heatherly, University of Southwestern Louisiana.

11. *Ultrafilters and* C_1 *fields*, by Brother Joseph Heisler, St. Edward's University.

12. *Relations between critical point theory and differential systems*, by W. M. Whyburn, Southern Methodist University and East Carolina University.

13. *Reducible linear differential operators*, by. D. H. Anderson, Southern Methodist University.

14. *Canonical forms for an almost linear partial differential equation of third order*, by Russell Cowan, Lamar State College of Technology.

15. *Solutions of second-order partial differential equations and conditions for existence of complementary functions*, by L. V. Robinson, Argyle, Texas.

16. *A class of polynomials generated by* $\exp (x/3)^3 \cdot d^n/dx^n \exp (-x/3)^3$, by T. M. Howle, Jr. and John Perryman, University of Texas at Arlington.

17. *A complex inversion formula for an integral transform with kernel* $\exp(-st^3)\ H_n(s, t)$, by Stephen West and John Perryman, University of Texas at Arlington.

18. *A linear integral transform with a simple kernel*, by W. W. Bolton and S. C. Crim, Lamar State College of Technology.

19. *Expansive flows*, by R. K. Williams, Southern Methodist University.

20. *A note on sequential convergence in compact Hausdorff spaces without perfect subsets*, by R. V. McPherson, University of Texas at Austin.

21. P_n *functions and n-convex sets*, by M. W. Jeter, Stillwater, Oklahoma.

22. *A generalization of a theorem of F. Riesz*, by F. N. Huggins, University of Texas at Arlington.

23. *An old soldier looks at academic freedom*, by J. W. Strain, Midwestern University.

24. *Testing unusual hypotheses about a contingency table*, by Ronald Harrist and Wanzer Drane, Southern Methodist University.

25. *Using the computer in courses in calculus and logic*, by Joe Wimbish and Dale Maness, Oklahoma College of Liberal Arts.

26. *On extrema at the boundary*, by A. R. Amir-Moez, Texas Tech. University.

27. *Some applications of nonstandard analysis to theory of differential equations*, by Vadim Komkov, Texas Tech University.

28. *Some mathematical foundations for interdisciplinary studies*, by Dale Maness and Joe Wimbish, Oklahoma College of Liberal Arts.

29. *On the problem of transfer of credits*, by Glen Mattingly, Sam Houston State University.

30. *Some modifications of Muller's method for finding roots of ill-conditioned polynomials*, by B. L. Turlington and Donna K. Dunaway, Southern Methodist University.

31. *Spherical programming: a new approach to convex programming*, by S. W. McGuire, Lamar State College of Technology.

32. *An exact partition of Wilk's chi-square*, by Ronald Harrist, Southern Methodist University.

J. C. Bradford, *Secretary-Treasurer*

MAY MEETING OF THE MICHIGAN SECTION

The annual meeting of the Michigan Section of the MAA was held at Western Michigan University, Kalamazoo, Michigan on May 7 and 8, 1971. There were approximately 180 people in attendance. This was the first time that the Michigan Section had used the two-day format, and the results were encouraging.

Professor Jay Folkert of Hope College, Chairman of the Section, presided at the business meeting. The Section approved a revision of its By-Laws, voted to reaffiliate with the Michigan Academy of Arts and Sciences, and passed a resolution requesting the National Organization to collect dues for the sections. Officers elected for the coming year were: M. T. Wechsler, Wayne State University, Chairman; D. J. Lewis, University of Michigan, Vice-Chairman; and Yousef Alavi, Western Michigan University, Secretary-Treasurer.

Invited addresses were given by Dr. Gorton Riethmiller, member of the State Board of Education, and Professor J. S. Frame, Michigan State University. The title of Professor Frame's talk was "Square Root Enumeration and Group Characters."

The program also featured four panel discussions and a live demonstration of a grade school class being taught by a member of the SEED Project. The make-up of the panels was as follows:

Panel on Curriculum Problems in the Two-year College. Lead-off speaker: Donald Ross, Washtenaw Community College. Responders: Lowell Stultz, Kalamazoo Valley Community College; Katherine Price, Highland Park Junior College; and William Lakey, Central Michigan University. Moderator: A. B. Clarke, Western Michigan University.

Panel on Computer Science and the Undergraduate Mathematics Program. Lead-off speaker: Sam Conte, Purdue University. Responders: Roy Jorgensen, Andrews University; Franklin Westervelt, Wayne State University; and James Powell, Western Michigan University. Moderator: J. S. Frame, Michigan State University.

Panel on Accreditation and Certification. Lead-off speaker: Billy Rhoades, Indiana University. Responders: Wade Ellis, University of Michigan; Raymond Spencer, Henry Ford Community College; Richard Vandevelde, Hope College; and Wilbur Walkoe, Jr., Grand Valley State College. Moderator: Lyle Mehlenbacher, University of Detroit.

Panel on the Role of Geometry in the Undergraduate Curriculum. Lead-off speaker: Nicholas Kazarinoff, University of Michigan. Responders: Robert Kane, University of Detroit; James McKay, Oakland University; and B. M. Stewart, Michigan State University. Moderator: L. M. Kelly, Michigan State University.

The following papers were presented:

1. *Still Another Elementary Proof that* $\Sigma 1/k^2=\pi^2/6$, by D. P. Giesy, Western Michigan University.
2. *The Lagrange Multiplier Rule*, by M. S. Skaff, University of Detroit.
3. *Paths in Street Networks*, by Edward Nordhaus, Michigan State University.
4. *A Geometric View of Power Series*, by George Piranian, University of Michigan.
5. *An Addition Theorem for Polynomial-Exponential Functions*, by P. M. Anselone, Oregon State University and Michigan State University.

H. T. SLABY, *Secretary-Treasurer*

MAY MEETING OF THE UPPER NEW YORK STATE SECTION

The Spring Meeting of the Upper New York State Section of the MAA was held at St. Lawrence University, Canton, N. Y., on May 8, 1971. There were 80 members and 18 guests in attendance.

The Fourth Annual Harry M. Gehman Invited Lecture, "Backward, Turn Backward," was given by Professor J. A. F. Randolph of the Rochester Institute of Technology, and Fayerweather Professor of Mathematics, Emeritus, University of Rochester.

Other papers on the program were:

1. *Integration by diagrams*, by D. S. Martin, State University College at Brockport.
2. *A new direction for Birkhoff's Problem 111*, by R. C. Shiflett, Wells College.
3. *Invariance theorems in probability*, by D. R. Beuerman, Queen's University.
4. *Seventeen wallpaper samples*, by D. Paine, Wells College and Cornell University.
5. *Informal vs. formal mathematics*, by R. G. van Meter, St. Lawrence University.

At the business meeting the following officers were elected: C. F. Stephens, State University College at Potsdam, Chairman; E. Hemmingsen, Syracuse University, Vice-Chairman; P. T. Schaefer, State University College at Geneseo, Secretary-Treasurer.

The Two-Year College Committee, consisting of L. A. Trivieri, Mohawk Valley Community College; C. A. Lathan, Monroe Community College; J. D. Vadney, Fulton-Montgomery Community College; P. W. Gilbert, Syracuse University; and M. F. Smiley, SUNY at Albany; presented its report which recommended that (1) a joint meeting with other organizations be held to discuss problems of articulation of the two-year colleges with four-year colleges of the section, (2) a vice-chairman for two-year college affairs be added to the executive committee of the section, (3) that two-year college faculty be more widely represented on CUPM panels and in the Visiting Mathematician Program, and (4) that the question of reciprocity agreements for joint memberships between MAA and other mathematics organizations be explored.

PAUL SCHAEFER, *Secretary-Treasurer*

JUNE MEETING OF THE PACIFIC NORTHWEST SECTION

The annual meeting of the Pacific Northwest Section of the MAA was held at Oregon State University, Corvallis, June 18 and 19, 1971 in conjunction with the Six Hundred Eighty-Sixth meeting of AMS. One hundred sixty-eight persons were in attendance, of which one hundred thirteen were members of the MAA.

At the business meeting the following officers were elected: Chairman, Maurice Kingston, University of Washington; First Vice-Chairman, Sheldon Rio, Southern Oregon College; Second Vice-Chairman, Edward James, Edmonds Community College; Secretary-Treasurer, James Calvert, University of Idaho.

The Bylaws of the Section were amended as suggested by the national organization to enable the Section to acquire tax-exempt status under the Internal Revenue Code. The Section also adopted a resolution urging the national organization to investigate with NCATE the possibility of having a person approved by the national MAA organization on every NCATE evaluation team and that, if necessary, the national organization underwrite the expenses of such representation. It is thought that the presence of an MAA representative on NCATE evaluation teams would help implement CUPM recommendations for the training of teachers.

It was announced that the Section prizes of \$30, \$25, and \$20 awarded annually to the top three scorers from the Section in the Putnam Competition were received, respectively, by Joe Buhler, Reed College, John Mallet-Paret, University of Alberta, and Kent Brothers, University of Victoria.

The program of the MAA portion of the meeting was as follows:

1. *Fourier analysis and the origins of set theory*, by Keith Yale, University of Montana.
2. *Shapes of the future*, by Victor Klee, University of Washington.
3. *Panel discussion: Should the standard freshman calculus course be taught with the aid of computers?* Moderator: Samuel Dunn, Seattle Pacific College; panelists: Donald Guthrie, Oregon State University; Billy Hobbs, Pasadena College; Roger Jay, Lane Community College; David Moursund, University of Oregon.
4. *Linear orders*, by Jack Robertson, Washington State University.
5. *Panel discussion: Teaching occupational math.* Moderator: Howard Zink, Lane Community College; panelists: Bud Cooke, Lane Community College; Ron Waite, Blue Mountain Community College; Frank Weeks, Mt. Hood Community College.
6. *Panel discussion: An instructional system for occupational and developmental students.* Moderator: Lawrence Mitchell, Blue Mountain Community College; panelists: Harold Hauser, Mickey McClendon, Ron Waite and Leon Severin, Blue Mountain Community College; Edward Wright, Linn-Benton Community College.
7. *Panel discussion: Using the audio-tutorial approach.* Moderator: Edward Wright, Linn-Benton Community College; panelists: Michael Greenwood, Clark College; Robert Main, Oregon College of Education; Mickey McClendon, Blue Mountain Community College.

E. A. MAIER, *Secretary*

AAAS SECTION ON MATHEMATICS CO-SPONSORED BY THE MAA AND NCTM

For the first time, the MAA will co-sponsor with NCTM a session on mathematics at the AAAS meeting. This jointly sponsored session will be held on Tuesday, December 28, 1971 in the Academy Room of the Bellevue-Stratford in Philadelphia.

The general purpose of this session is to examine both philosophy and actual practice in the area of the teaching of applied mathematics. The main emphasis will be on a number of highly interesting and successful current projects, ranging from elementary through secondary schools into the colleges. Of particular interest will be a number of attempts to bring real applications of the mathematical sciences into the classroom.

The session has been arranged by Dr. H. O. Pollak, Director, Mathematics and Statistics Research Center, Bell Telephone Laboratories, Murray Hill, New Jersey. The program, entitled "The Relation Between the Applications of Mathematics and Teaching of Mathematics," will be the following:

TUESDAY, DECEMBER 28 Bellevue-Stratford, Academy Room
9:00 A.M. Chairman: HENRY O. POLLAK

Correlating Science and Mathematics in the Elementary School: The USMES Project
WILLIAM M. FITZGERALD (*Professor, Mathematics Education, Michigan State University, East Lansing, Michigan*).

Outline of a Problem Oriented, Applications Oriented, and Computer Oriented High School Mathematics Course
ARTHUR ENGEL (*Professor, Ludwigsburg College, Ludwigsburg, Germany*).

Some Experiences in Teaching Mathematical Modelling to Undergraduates
MAYNARD D. THOMPSON (*Professor of Mathematics, Indiana University, Bloomington, Indiana*).

2:00 P.M. Chairman: HENRY O. POLLAK

Real Statistics in the Secondary School: Activities of the ASA-NCTM Joint Committee
WILLIAM H. KRUSKAL (*Professor, Department of Statistics, University of Chicago, Chicago, Illinois*).

The Man-Made World
JOHN G. TRUXAL (*Academic Vice-President, Polytechnic Institute of Brooklyn, Brooklyn, New York*) and EMIL J. PIEL (*Executive Director, ECCP*).

The Undergraduate Mathematics Curriculum and Applied Mathematics: The Work of CUPM
MAYNARD D. THOMPSON.

Some General Comments on the Teaching of Applications of Mathematics
HENRY O. POLLAK.

ACKNOWLEDGMENT

The editors wish to thank the following mathematicians who have refereed manuscripts for Volume 78: T. M. Apostol, R. G. Aroub, R. A. Askey, M. Barr, R. G. Bartle, E. F. Beckenbach, L. W. Beineke, R. E. Bellman, S. K. Berberian, L. D. Berkowitz, R. H. Bing, E. A. Bishop, D. Blackwell, S. L. Bloom, R. P. Boas, F. Brauer, L. Breiman, A. M. Bruckner, P. B. Burcham, H. Busemann, F. P. Callahan, L. Carlitz, G. D. Chakerian, G. T. Chartrand, H. Chernoff, L. N. Childs, P. T. Church, E. A. Coddington, G. E. Collins, H. L. Croft, R. B. Darst, C.-W. R. de Boor, C. Dombrowski, J. L. Doob, M. P. Drazin, J. Dugundji, S. Eilenberg, H. Federer, P. A. Fillmore, H. Flanders, L. Flatto, W. H. Fleming, G. E. Forsythe, D. J. Foulis, S. P. Franklin, R. W. Freese, P. J. Freyd, A. Friedman, S. A. Gaal, R. Gilmer, C. Goffman, J. Goldman, M. Golomb, W. H. Gottschalk, F. A. Graybill, N. Grossman, E. Grosswald, B. Grunbaum, C. P. Gupta, M. Guterman, P. Hagis, Jr., P. Hagis, H. Halberstam, M. Hall, Jr., P. R. Halmos, F. Harary, P. Hartman, W. J. Heinzer, I. N. Herstein, E. Hewitt, P. J. Hilton, I. I. Hirschman, Jr., A. J. Hoffman, N. Jacobson, M. Jerison, F. John, J. H. Jordan, W. M. Kantor, S. Kaplan, N. D. Kazarinoff, M. R. Kirch, V. L. Klee, Jr., D. J. Kleitman, M. Knopp, R. R. Korfhage, J. Lambek, S. Lang, P. Lax, E. L. Lehmann, D. H. Lehmer, W. Leighton, W. J. LeVeque, N. Levinson, J. Lipman, G. G. Lorentz, R. E. Lynch, R. C. Lyndon, G. R. MacLane, S. MacLane, M. Marcus, A. P. Mattuck, K. O. May, E. J. McShane, G. H. Meisters, E. A. Michael, J. B. Miles, B. M. Mitchell, B. Mond, D. S. Moore, R. A. Morris, L. Muenz, T. W. Mullikin, M. E. Munroe, I. Namioka, A. Nerode, M. F. Neuts, A. Nijenhuis, L. Nirenberg, I. Niven, J. A. Nohel, J. M. H. Olmsted, G. Orzech, M. Orzech, L. J. Paige, G. Papanicolau, J. L. Paul, D. Pedoe, S. Perlis, R. R. Phelps, W. S. Piper, J. M. Plotkin, H. Popp, W. Prenowitz, W. V. O. Quine, R. Rado, A. Ralston, I. Reiner, C. E. Richart, T. J. Rivlin, J. L. Roberts, D. H. Root, A. Rosenberg, G.-C. Rota, J. J. Rotman, H. Rubin, J. E. Rubin, M. E. Rudin, W. Rudin, B. Ryan, H. J. Ryser, S. M. Samuels, D. W. Sasser, M. M. Schacher, E. V. Schenkman, A. Schild, W. R. Scott, G. J. Simmons, S. Smale, E. Snapper, L. J. Snell, R. I. Soare, F. L. Spitzer, L. A. Steen, R. Steinberg, A. K. Steiner, B. M. Stewart, K. Stolarsky, A. H. Stone, M. V. Subbarao, O. Taussky, A. E. Taylor, W. T. Tutte, S. L. Ulam, W. R. Utz, F. A. Valentine, D. H. Van Osdol, J. S. P. Wang, D. Waterman, C. E. Watts, S. Weingram, G. Weiss, J. G. Wendel, A. Wilansky, J. W. Yackel, P. R. Young, H. J. Zassenhaus, D. Zelinsky, R. E. Zink.

CALENDAR OF FUTURE MEETINGS

Fifty-fifth Annual Meeting, Las Vegas, Nevada, January 19–21, 1972.

Fifty-third Summer Meeting, Dartmouth College, Hanover, New Hampshire, August 28–30, 1972.

The following is a list of the Sections of the Association with dates of future meetings so far as they have been reported to the Editorial Director.

ALLEGHENY MOUNTAIN, Pennsylvania State University, Altoona, May 5–6, 1972.

FLORIDA, Central Florida Junior College, Ocala, March 17–18, 1972.

ILLINOIS, Lake Forest College, Lake Forest, May 12–13, 1972.

INDIANA

IOWA, University of Iowa, Iowa City, April 28, 1972.

KANSAS, Washburn University, Topeka, March 24–25, 1972.

KENTUCKY, Georgetown University, Georgetown, Spring 1972.

LOUISIANA-MISSISSIPPI, Millsaps College, Jackson, Mississippi, February 18–19, 1972.

MARYLAND-DISTRICT OF COLUMBIA-VIRGINIA

METROPOLITAN NEW YORK

MICHIGAN, Oakland University, Rochester, May 5–6, 1972.

MISSOURI, Stephens College, Columbia, May 5–6, 1972.

NEBRASKA, University of Nebraska at Omaha, Omaha, April 21–22, 1972.

NEW JERSEY

NORTH CENTRAL

NORTHEASTERN

NORTHERN CALIFORNIA, California State College at Hayward, Hayward, February 5, 1972.

OHIO, Wittenberg University, Springfield, April 28–29, 1972.

OKLAHOMA-ARKANSAS, State College of Arkansas, Conway, Arkansas, March 10–11, 1972.

PACIFIC NORTHWEST, University of Washington, Seattle, Washington, June 16–17, 1972.

PHILADELPHIA

ROCKY MOUNTAIN, Southern Colorado State College, Pueblo, May 5–6, 1972.

SOUTHEASTERN, Samford University, Birmingham, Alabama, March 24–25, 1972.

SOUTHERN CALIFORNIA, California Institute of Technology, Pasadena, March 11, 1972.

SOUTHWESTERN, University of New Mexico, Albuquerque, Spring 1972.

TEXAS, Southwest Texas State University, San Marcos, April 1972.

UPPER NEW YORK STATE

WISCONSIN, Wisconsin State University, Stevens Point, April 28–29, 1972.

FUTURE MEETINGS OF OTHER ORGANIZATIONS

AMERICAN ASSOCIATION FOR THE ADVANCEMENT OF SCIENCE, Washington, D. C., December 26–30, 1972.

AMERICAN MATHEMATICAL SOCIETY, Las Vegas, Nevada, January 17–20, 1972.

AMERICAN SOCIETY FOR ENGINEERING EDUCATION

ASSOCIATION FOR COMPUTING MACHINERY, Boston, Massachusetts, August 14–16, 1972.

ASSOCIATION FOR SYMBOLIC LOGIC, Statler Hilton Hotel, New York City, December 27–28, 1971.

CENTRAL ASSOCIATION OF SCIENCE AND MATHEMATICS TEACHERS, Chicago, Illinois, November 16–18, 1972.

FIBONACCI ASSOCIATION

INSTITUTE OF MATHEMATICAL STATISTICS

MU ALPHA THETA, Chicago, Illinois, April 18, 1972.

NATIONAL COUNCIL OF TEACHERS OF MATHEMATICS, Chicago, Illinois, April 16–19, 1972.

OPERATIONS RESEARCH SOCIETY OF AMERICA, Jung Hotel, New Orleans, April 26–28, 1972.

PI MU EPSILON, Dartmouth College, Hanover, August 29–30, 1972.

SOCIETY FOR INDUSTRIAL AND APPLIED MATHEMATICS, Benjamin Franklin Hotel, Philadelphia, June 12–14, 1972 (20th Anniversary Celebration).

GEORGE BANTA COMPANY, INC., MENASHA, WISCONSIN